AD-A059 876 CARNEGIE-MELLON UNIV PITTSBURGH PA MANAGEMENT SCIENC--ETC F/G 12/2

DISJUNCTIVE PROGRAMMING.(U)

AUG 77 E BALAS N00014-75-C-0621

UNCLASSIFIED MSRR-415 NL

1 OF 1

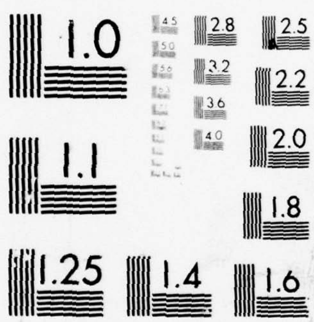
AD A059876

END

DATE FILMED

12-78

DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

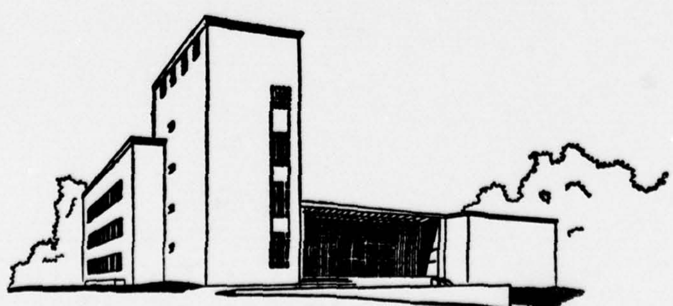
12

Carnegie-Mellon University

PITTSBURGH, PENNSYLVANIA 15213

GRADUATE SCHOOL OF INDUSTRIAL ADMINISTRATION

WILLIAM LARIMER MELLON, FOUNDER



78 10 03 025

Management Sciences Research Report No. 415

MSRR-

DISJUNCTIVE PROGRAMMING

by

Egon Balas

Aug 77

83 p.

Lecture Notes
for
Discrete Optimization 1977
Vancouver, British Columbia
August 8-12, 1977

This report was prepared as part of the activities of the Management Sciences Research Group, Carnegie-Mellon University, under Grant MPS73-08534 A02 of the National Science Foundation and Contract N00014-75-C-0621 NR 047-048 with the U.S. Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

NSF-MPS73-08534

Management Science Research Group
Graduate School of Industrial Administration
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

Abstract

This paper reviews some recent developments in the convex analysis approach to integer programming. These developments are based on viewing integer programs as disjunctive programs, i.e., linear programs with disjunctive constraints, an approach which seems to be fruitful for 0-1 programming both theoretically and practically. On the theoretical side, it provides structural characterizations which offer new insights. On the practical side, it produces a variety of cutting planes with desirable properties and offers new ways of combining cutting planes with enumerative techniques.

Table of Contents

DISJUNCTIVE PROGRAMMING

by

Egon Balas

1. Introduction

This paper reviews some recent developments in the convex analysis approach to integer programming. A product of the last four years, these developments are based on viewing integer programs as disjunctive programs, i.e., linear programs with disjunctive constraints. Apart from the fact that this is the most natural and straightforward way of stating many problems involving logical conditions (dichotomies, implications, etc.), the disjunctive programming approach seems to be fruitful for zero-one programming both theoretically and practically. On the theoretical side, it provides neat structural characterizations which offer new insights. On the practical side, it produces a variety of cutting planes with desirable properties, and offers several ways of combining cutting planes with branch and bound.

The line of research that has led to the disjunctive programming approach originates in the work on intersection or convexity cuts by Young [40], Balas [2], Glover [23], Owen [36] and others (see also [13], [24], [42]). This geometrically motivated work can be described in terms of intersecting the edges of the cone originating at the linear programming optimum $\bar{x}$ with the boundary of some convex set S, whose interior contains $\bar{x}$ but no feasible integer points, and using the hyperplane defined by the intersection points as a cutting plane. An early forerunner of this kind of approach was a paper by Hoang Tuy [29].

In the early 70's, research on intersection or convexity cuts was pursued in two main directions. One, typified by [23], [18], [4], was aimed at obtaining stronger cuts by including into S some explicitly or implicitly enumerated feasible integer points. The other, initiated by [3], brought into play polarity, in particular outer polars (i.e., polars of the feasible set, scaled so as to contain all feasible 0-1 points in their boundary), and related concepts of convex analysis, like maximal convex extensions, support and gauge functions, etc. (see also [5], [15], [19]). Besides cutting planes, this second direction has also produced (see [5], [6]) a "constraint activating" method (computationally untested to date) based on the idea of "burying" the feasible set into the outer polar (without using cuts), by activating the problem constraints one by one, as needed. This research has yielded certain insights and produced reasonably strong cutting planes; but those procedures that were implemented (and, by the way, very few were) turned out to be computationally too expensive to be practically useful.

In 1973 Glover [25] discovered that intersection cuts derived from a convex polyhedron S can be strengthened by rotating the facets of S in certain ways, a procedure he called polyhedral annexation. This was an important step toward the development of the techniques discussed in this paper. The same conclusions were reached independently (and concomitantly) in a somewhat different context by Balas [7]. The new context was given by the recognition that intersection cuts could be viewed as derived from a disjunction. Indeed, requiring that no feasible integer point be contained in the interior of S, is the same as requiring every feasible integer point to satisfy at least one of the inequalities whose complements define the supporting halfspaces of S. This seemingly innocuous change of perspective proved to be extremely fruitful. For one thing it led naturally and immediately to the consideration of disjunctive programs in their generality [8], [9], and hence to a characterization of all the valid inequalities for an integer program. By the same token, it offered the new possibility of generating cuts specially tailored for problems with a given structure. Besides, it offered a unified theoretical perspective on cutting planes and enumeration, as well as practical ways of combining the two approaches. Finally, it vastly simplified the proofs of many earlier results and opened the way to the subsequent developments to be discussed below.

Besides the above antecedents of the disjunctive programming approach there have been a few other papers concerned with linear (or nonlinear) programs with disjunctive constraints [26], [27]. The paper by Owen [36] deserves special mention, as the first occurrence of a cut with coefficients of different signs. However, these efforts were focused on special cases.

The main conceptual tool used in studying the structural properties of disjunctive programs is polarity. Besides the classical polar sets, we use a natural generalization of the latter which we call reverse polar sets [10]. This connects our research to the work of Fulkerson [20], [21], [22], whose blocking and antiblocking polyhedra are close relatives of the polar and reverse polar sets. There are also connections with the work of Tind [39] and Araoz [1]. One crucial difference in the way polars and reverse polars of a polyhedron are used in our work, versus that of the above mentioned authors, is the fact that we "dualize" the reverse polar $S^{\#}$ of a (disjunctive) set S, by representing it in terms of the inequalities (of the disjunction) defining S, rather than in terms of the points of S. It is precisely this element which leads to a linear programming characterization of the convex hull of feasible points of a disjunctive program.

Except for the specific applications described in sections 7 and 8, which are new, the results reviewed here are from [8], [9], [10], [14].

We tried to make this review self-contained, giving complete proofs for most of the statements. Most of the results are illustrated on numerical examples. For further details and related developments the reader is referred to the above papers, as well as those of Glover [25], Jeroslow [32], and Balas [11]. Related theoretical developments are also to be found in Jeroslow [30], [31], Blair and Jeroslow [17], Zemel [41], while some related computational procedures are discussed in Balas [12].

This paper is organized as follows. Section 2 introduces some basic concepts and terminology, and discusses ways of formulating disjunctive programs (DP). Section 3 contains the basic characterization of the family of valid inequalities for a DP, and discusses some of the implications of this general principle for deriving cutting planes. Section 4 extends the duality theorem of linear programming to DP. Section 5 discusses the basic properties of reverse polars and uses them to characterize the convex hull of a disjunctive set (the feasible set of a DP). It shows how the facets of the convex hull of the feasible set of a DP in n variables, defined by a disjunction with q terms, can be obtained from a linear program with $q(n + 1)$ constraints.

In section 6 we address the intriguing question of whether the convex hull of a disjunctive set can be generated sequentially, by imposing one by one the disjunctions of the conjunctive normal form, and producing at each step the convex hull of a disjunctive set with one elementary disjunction. We answer the question in the negative for the case of a general DP or a general integer program, but in the positive for a class of DP called facial, which subsumes the general (pure or mixed) zero-one program, the linear complementarity problem, the nonconvex (linearly constrained) quadratic program, etc.

The first 6 sections treat the integrality constraints of an (otherwise) linear program as disjunctions. When it comes to generating cutting planes from a particular noninteger simplex tableau, the disjunctions that can be used effectively are those involving the basic variables. Section 7 discusses a principle for strengthening cutting planes derived from a disjunction, by using the integrality conditions (if any) on the nonbasic variables. Section 8 illustrates on the case of multiple choice constraints how the procedures of sections 3 and 7 can take advantage of problem structure. Section 9 discusses some ways in which disjunctive programming can be used to combine branch and bound with cutting planes. In particular, if LP_k, $k \in Q$, are the subproblems associated with the active nodes of the search tree at a given stage of a branch and bound process applied to a mixed integer program P, it is shown that a cut can be derived from the cost rows of the simplex tableaux associated with the problems LP_k, which provides a bound on the value of P, often considerably better than the one available from the branch and bound process. Finally, section 10 deals with techniques for deriving disjunctions from conditional bounds. Cutting planes obtained from such disjunctions have been used with good results on large sparse set covering problems.

2. Linear Programs with Logical Conditions

By disjunctive programming we mean linear programming with disjunctive constraints. Integer programs (pure and mixed) and a host of other nonconvex programming problems (the linear complementarity problem, the general quadratic program, separable nonlinear programs, bimatrix games, etc.) can be stated as linear programs with logical conditions. In the present context "logical conditions" means statements about linear inequalities, involving the operations "and" ($\wedge$, conjunction--sometimes denoted by juxtaposition), "or" ($\vee$, disjunction), "complement of" ($\neg$, negation). The operation "if...then" ($\Rightarrow$, implication) is known to be equivalent to a disjunction. The negation of a conjunctive set of inequalities is a disjunction whose terms are the same inequalities. The operation of conjunction applied to linear inequalities gives rise to (convex) polyhedral sets. The disjunctions are thus the crucial elements of a logical condition (the ones that make the constraint set nonconvex), and that is why we call this type of problem a <u>disjunctive program</u>.

A disjunctive program (DP) is then a problem of the form

$$\min \{cx \mid Ax \geq a_0, \ x \geq 0, \ x \in L\} \quad .$$

where A is an $m \times n$ matrix, a_0 is an m-vector, and L is a set of logical conditions. Since the latter can be expressed in many ways, there are many different forms in which a DP can be stated. Two of them are fundamental.

The constraint set of a DP (and the DP itself) is said to be in <u>disjunctive normal form</u> if it is defined by a disjunction whose terms do not contain further disjunctions; and in <u>conjunctive normal form</u>, if it is defined by a conjunction whose terms do not contain further conjunctions.

The disjunctive normal form is then

$$\bigvee_{h\in Q} \begin{pmatrix} A^h x \geq a_0^h \\ x \geq 0 \end{pmatrix} \tag{2.1}$$

while the conjunctive normal form is

$$\begin{aligned} &Ax \geq a_0 \\ &x \geq 0 \\ &\bigvee_{i\in Q_j} (d^i x \geq d_{i0}), \quad j \in S \end{aligned} \tag{2.2}$$

or, alternatively,

$$\begin{pmatrix} Ax \geq a_0 \\ x \geq 0 \end{pmatrix} \wedge \left[\bigvee_{i\in Q_1} (d^i x \geq d_{i0})\right] \wedge \ldots \wedge \left[\bigvee_{i\in Q_{|S|}} (d^i x \geq d_{i0})\right]. \tag{2.2'}$$

Here each d^i is an n-vector and each d_{i0} a scalar, while the sets Q and Q_j, $j \in S$, may or may not be finite. The connection between the two forms is that each term of the disjunction (2.1) has, besides the $m+n$ inequalities of the system $Ax \geq a_0$, $x \geq 0$, precisely one inequality $d^i x \geq d_{i0}$, $i \in Q_j$, from each disjunction $j \in S$ of (2.2), and that all distinct systems $A^h x \geq a_0^h$, $x \geq 0$ with this property are present among the terms of (2.1); so that, if Q (and hence each Q_j, $j \in S$) is finite, then $|Q| = |\prod_{j\in S} Q_j|$, where π stands for cartesian product. Since the operations $\wedge$ and $\vee$ are distributive with respect to each other [i.e., if A, B, C are inequalities, $A \wedge (B \vee C) = AB \vee AC$, and $A \vee (BC) = (A \vee B) \wedge (A \vee C)$], any logical condition involving these operations can be brought to any of the two fundamental forms, and each of the latter can be obtained from the other one.

We illustrate the meaning of these two forms on the case when the DP is a zero-one program in n variables. Then the disjunctive normal form (2.1) is

$$\bigvee_{h\in Q} (Ax \geq 0,\ x \geq 0,\ x = x^h)$$

where $x^1,\ldots,x^{|Q|}$ is the set of all 0-1 points, and $|Q| = 2^n$; whereas the conjunctive normal form is

$$Ax \geq 0,\ x \geq 0,\ (x_j = 0) \vee (x_j = 1),\ j = 1,\ldots,n.$$

Once the inequalities occurring in the conjunctions and/or disjunctions of a DP are given, the disjunctive and conjunctive normal forms are unique. It is a fact of crucial practical importance, however, that the inequalities expressing the conditions of a given problem can be chosen in more than one way. For instance, the constraint set

$$\begin{aligned} 3x_1 + x_2 - 2x_3 + x_4 &\leq 1 \\ x_1 + x_2 + x_3 + x_4 &\leq 1 \\ x_j = 0 \text{ or } 1,\ j = 1,\ldots,4, \end{aligned}$$

when put in disjunctive normal form, becomes a disjunction with $2^4 = 16$ terms; but the same constraint set can also be expressed as

$$3x_1 + x_2 - 2x_3 + x_4 \leq 1$$

$$\bigvee_{i=1}^{4} \begin{pmatrix} x_i = 1 \\ x_j = 0,\ j \neq i \end{pmatrix} \vee (x_j = 0,\ \forall j)$$

which gives rise to a disjunction with only 5 terms.

3. The Basic Principle of Disjunctive Programming

A constraint B is said to be a consequence of, or implied by, a constraint A, if every x that satisfies A also satisfies B. We are interested in the family of inequalities implied by the constraint set of a general disjunctive program (DP). All valid cutting planes for a DP belong of course to this family. On the other hand, the set of points satisfying all the inequalities in the family is precisely the convex hull of the set of feasible solutions to the DP. A characterization of this family is given in the next theorem, which is an easy but important generalization of a classical result. Let $x \in R^n$, $\alpha \in R^n$, $\alpha_0 \in R$, $A^h \in R^{m_h \times n}$, $a_0^h \in R^{m_h}$, $h \in Q$ (not necessarily finite) and let a_j^h be the j-th column of A^h, $h \in Q$, $j \in N = \{1,\dots,n\}$.

Theorem 3.1. The inequality $\alpha x \geq \alpha_0$ is a consequence of the constraint

$$(3.1) \qquad \bigvee_{h \in Q} \begin{pmatrix} A^h x \geq a_0^h \\ x \geq 0 \end{pmatrix}$$

if and only if there exists a set of $\theta^h \in R^{m_h}$, $\theta^h \geq 0$, $h \in Q^*$, satisfying

$$(3.2) \qquad \alpha \geq \theta^h A^h \quad \text{and} \quad \alpha_0 \leq \theta^h a_0^h \,, \quad \forall h \in Q^*,$$

where Q^* is the set of those $h \in Q$ such that the system $A^h x \geq a_0$, $x \geq 0$ is consistent.

Proof. $\alpha x \geq \alpha_0$ is a consequence of (3.1) if and only if it is a consequence of each term $h \in Q^*$ of (3.1). But according to a classical result on linear inequalities (see, for instance, Theorem 1.4.4 of [37], or Theorem 22.3 of [36]), this is the case if and only if the conditions stated in the theorem hold. ||

Remark 3.1.1. If the i-th inequality of a system $h \in Q^*$ of (3.1) is replaced by an equation, the i-th component of θ^h is to be made unconstrained.

If the variable x_j in (3.1) is allowed to be unconstrained, the j-th inequality of each system $\alpha \geq \theta^h A^h$, $h \in Q^*$, is to be replaced by the corresponding equation in the "if" part of the statement.

With these changes, Theorem 3.1 remains true.

An alternative way of stating (3.2) is

$$(3.3) \qquad \begin{aligned} \alpha_j &\geq \sup_{h \in Q^*} \theta^h a_j^h, \quad j \in N \\ \alpha_0 &\leq \inf_{h \in Q^*} \theta^h a_0^h. \end{aligned}$$

Since $Q^* \subseteq Q$, the if part of the Theorem remains of course valid if Q^* is replaced by Q.

Since (3.3) defines <u>all</u> the valid inequalities for (3.1), every valid cutting plane for a disjunctive program can be obtained from (3.3) by choosing suitable multipliers θ_i^h. If we think of (3.1) as being expressed in terms of the nonbasic variables of a basic optimal solution to the linear program associated with a DP, then a valid inequality $\alpha x \geq \alpha_0$ cuts off the optimal linear programming solution (corresponding to $x_j = 0$, $j \in N$) if and only if $\alpha_0 > 0$; hence α_0 will have to be fixed at a positive value. Inequalities with $\alpha_0 \leq 0$ may still cut off parts of the linear programming feasible set, but not the optimal solution $x = 0$.

The special case when each system $A^h x \geq a_0^h$, $h \in Q$, consists of the single inequality $a^h x \geq a_{h0}$ (a^h a vector, a_{h0} a positive scalar) deserves special mention. In this case, choosing multipliers $\theta^h = 1/a_{h0}$, $h \in Q$, we obtain the inequality

$$(3.4) \qquad \sum_{j \in J} \left(\max_{h \in Q} a_j^h / a_{h0}\right) x_j \geq 1 ,$$

which (for Q finite) is Owen's cut [36]. It can also be viewed as a slightly improved version of the intersection cut from the convex set

$$S = \{x \mid a^h x \leq a_{h0},\ h \in Q\},$$

which has the same coefficients as (3.4) except for those (if any) $j \in J$ such that $a_j^h < 0$, $\forall h \in Q$. For the latter, the intersection cut from S has zero coefficients whereas the corresponding coefficients of (3.4) are negative.

Whenever all the coefficients of (3.4) are positive (in terms of intersection cuts, this corresponds to the case when S is bounded), (3.4) is the strongest inequality implied by the disjunction $\bigvee_{h \in Q} (a^h x \geq a_{h0})$; in the presence of negative coefficients, however, (3.4) can sometimes be further strengthened.

Due to the generality of the family of inequalities defined by (3.3), not only can the earlier cuts of the literature be easily recovered by an appropriate choice of the multipliers θ^h (see [8] for details), but putting them in the form (3.3) indicates, by the same token, ways in which they can be strengthened by appropriate changes in the multipliers.

A significant new feature of the cutting planes defined by (3.3) consists in the fact that they can have coefficients of different signs. The classical cutting planes, as well as the early intersection/convexity cuts and the group theoretic cutting planes (including those corresponding to facets of the corner polyhedron), are all restricted to positive coefficients (when stated in the form $\geq$, in terms of the nonbasic variables of the tableau from which they were derived). This important limitation, which tends to produce degeneracy in dual cutting plane algorithms, can often be overcome in the case of the cuts obtained from (3.3) by an appropriate choice of multipliers.

Another important feature of the principle expressed in Theorem 3.1 for generating cutting planes is the fact that in formulating a given integer program as a disjunctive program, one can take advantage of any particular structure the problem may have. In section 7 we will illustrate this on some frequently occurring structures. We finish this section by an example of a cut for a general mixed integer program.

<u>Example 3.1</u>. Consider the mixed integer program whose constraint set is

$$x_1 = .2 + .4(-x_3) + 1.3(-x_4) - .01(-x_5) + .07(-x_6)$$
$$x_2 = .9 - .3(-x_3) + .4(-x_4) - .04(-x_5) + .1\,(-x_6)$$
$$x_j \geq 0,\ j = 1,\dots,6,\ x_j \text{ integer},\ j = 1,\dots,4.$$

This problem is taken from Johnson's paper [35], which also lists six cutting planes derived from the extreme valid inequalities for the associated group problem:

$$\begin{aligned}
.75\,x_3 + .875x_4 + .0125x_5 + .35\,x_6 &\geq 1\\
.778x_3 + .444x_4 + .40\,x_5 + .111x_6 &\geq 1\\
.333x_3 + .667x_4 + .033x_5 + .35\,x_6 &\geq 1\\
.50\,x_3 + x_4 + .40\,x_5 + .25\,x_6 &\geq 1\\
.444x_3 + .333x_4 + .055x_5 + .478x_6 &\geq 1\\
.394x_3 + .636x_4 + .346x_5 + .155x_6 &\geq 1\,.
\end{aligned}$$

The first two of these inequalities are the mixed-integer Gomory cuts derived from the row of x_1 and x_2 respectively. To show how they can be improved, we first derive them as they are. To do this, for a row of the form

$$x_i = a_{i0} + \sum_{j\in J} a_{ij}(-x_j)\ ,$$

with x_j integer-constrained for $j \in J_1$, continuous for $j \in J_2$, one defines $f_{ij} = a_{ij} - [a_{ij}]$, $j \in J \cup \{0\}$, and $\varphi_{i0} = f_{i0}$,

$$\varphi_{ij} = \begin{cases} f_{ij} & j \in J_1^+ = \{j \in J_1 \mid f_{i0} \geq f_{ij}\} \\ f_{ij} - 1 & j \in J_1^- = \{j \in J_1 \mid f_{i0} < f_{ij}\} \\ a_{ij} & j \in J_2 \,. \end{cases}$$

Then every x which satisfies the above equation and the integrality constraints on x_j, $j \in J_1 \cup \{i\}$, also satisfies the condition

$$y_i = \varphi_{i0} + \sum_{j \in J} \varphi_{ij}(-x_j) \,, \qquad y_i \text{ integer} \,.$$

For the two equations of the example, the resulting conditions are

$$y_1 = .2 - .6(-x_3) - .7(-x_4) - .01(-x_5) + .07(-x_6), \quad y_1 \text{ integer}$$
$$y_2 = .9 + .7(-x_3) + .4(-x_4) - .04(-x_5) + .1\ (-x_6), \quad y_2 \text{ integer.}$$

Since each y_i is integer-constrained, they have to satisfy the disjunction $y_i \leq 0 \vee y_i \geq 1$. Applying the above theorem with multipliers $\theta_i = 1/a_{i0}$ then gives the cut (3.4) which in the two cases $i = 1,2$ is

$$\frac{.6}{.8} x_3 + \frac{.7}{.8} x_4 + \frac{.01}{.8} x_5 + \frac{.07}{.2} x_6 \geq 1$$

and

$$\frac{.7}{.9} x_3 + \frac{.4}{.9} x_4 + \frac{.04}{.1} x_5 + \frac{.1}{.9} x_6 \geq 1 \,.$$

These are precisely the first two inequalities of the above list.

Since all cuts discussed here are stated in the form ≥ 1, the smaller the j-th coefficient, the stronger is the cut in the direction j. We would thus like to reduce the size of the coefficients as much as possible.

Now suppose that instead of $y_1 \leq 0 \vee y_1 \geq 1$, we use the disjunction

$$\{y_1 \leq 0\} \vee \begin{Bmatrix} y_1 \geq 1 \\ x_1 \geq 0 \end{Bmatrix}$$

which of course is also satisfied by every feasible x.

Then, applying Theorem 3.1 with multipliers 5, 5 and 15 for $y_1 \leq 0$, $y_1 \geq 1$ and $x_1 \geq 0$ respectively, we obtain the cut whose coefficients are

$$\max \left\{ \frac{5 \times (-.6)}{5 \times .2}, \frac{5 \times .6 + 15 \times (-.4)}{5 \times .8 + 15 \times (-.2)} \right\} = -3$$

$$\max \left\{ \frac{5 \times (-.7)}{5 \times .2}, \frac{5 \times .7 + 15 \times (-1.3)}{5 \times .8 + 15 \times (-.2)} \right\} = -3.5$$

$$\max \left\{ \frac{5 \times (-.01)}{5 \times .2}, \frac{5 \times .01 + 15 \times .01}{5 \times .8 + 15 \times (-.2)} \right\} = .2$$

$$\max \left\{ \frac{5 \times .07}{5 \times .2}, \frac{5 \times (-.07) + 15 \times (-.07)}{5 \times .8 + 15 \times (-.2)} \right\} = .35 ,$$

that is

$$-3x_3 - 3.5x_4 + .2x_5 + .35x_6 \geq 1 .$$

The sum of coefficients on the left hand side has been reduced from 1.9875 to -5.95.

Similarly, for the second cut, if instead of $y_2 \leq 0 \vee y_2 \geq 1$ we use the disjunction

$$\begin{Bmatrix} y_2 \leq 0 \\ x_1 \geq 0 \end{Bmatrix} \vee \{y_2 \geq 1\} ,$$

with multipliers 10, 40 and 10 for $y_2 \leq 0$, $x_1 \geq 0$ and $y_2 \geq 1$ respectively, we obtain the cut

$$-7x_3 - 4x_4 + .4x_5 - x_6 \geq 1 .$$

Here the sum of left hand side coefficients has been reduced from 1.733 to -11.6.

4. Duality

In this section we state a duality theorem for disjunctive programs, which generalizes to this class of problems the duality theorem of linear programming.

Consider the disjunctive program

$$z_o = \min cx$$

$$\text{(P)} \qquad \bigvee_{h\in Q} \left\{ \begin{array}{c} A^h x \geq b^h \\ x \geq 0 \end{array} \right\},$$

where A^h is a matrix and b^h a vector, $\forall\, h\in Q$.

We define the dual of (P) to be the problem

$$w_o = \max w$$

$$\text{(D)} \qquad \bigwedge_{h\in Q} \left\{ \begin{array}{r} w - u^h b^h \leq 0 \\ u^h A^h \leq c \\ u^h \geq 0 \end{array} \right\}$$

The constraint set of (D) requires each u^h, $h\in Q$, to satisfy the corresponding bracketed system, and w to satisfy each of them.

Let

$$X_h = \{x \mid A^h x \geq b^h,\ x \geq 0\},\quad \bar{X}_h = \{x \mid A^h x \geq 0,\ x \geq 0\};$$

$$U_h = \{u^h \mid u^h A^h \leq c,\ u^h \geq 0\},\quad \bar{U}_h = \{u^h \mid u^h A^h \leq 0,\ u^h \geq 0\}.$$

Further, let

$$Q^* = \{h\in Q \mid X_h \neq \emptyset\} \quad , \quad Q^{**} = \{h\in Q \mid U_h \neq \emptyset\}.$$

We will assume the following

<u>Regularity condition</u>:

$$(Q^* \neq \emptyset,\ Q\backslash Q^{**} \neq \emptyset) \Rightarrow Q^*\backslash Q^{**} \neq \emptyset;$$

i.e., if (P) is feasible and (D) is infeasible, then there exists $h\epsilon Q$ such that $X_h \neq \emptyset$, $U_h = \emptyset$.

<u>Theorem 4.1</u>. Assume that (P) and (D) satisfy the regularity condition. Then exactly one of the following two situations holds.

1. Both problems are feasible; each has an optimal solution and $z_o = w_o$.

2. One of the problems is infeasible; the other one either is infeasible or has no finite optimum.

<u>Proof</u>. (i) Assume that both (P) and (D) are feasible. If (P) has no finite minimum, then there exists $h\epsilon Q$ such that $\bar{X}_h \neq \emptyset$ and $\bar{x} \in \bar{X}_h$ such that $c\bar{x} < 0$. But then $U_h = \emptyset$, i.e., (D) is infeasible; a contradiction.

Thus (P) has an optimal solution, say $\bar{x}$. Then the inequality $cx \geq z_o$ is a consequence of the constraint set of (P); i.e., $x \in X_h$ implies $cx \geq z_o$, $\forall\ h\epsilon Q$. But then for all $h\epsilon Q^*$, there exists $u^h \epsilon U_h$ such that $u^h b^h \geq z_o$. Further, since (D) is feasible, for each $h\epsilon Q\backslash Q^*$ there exists $\hat{u}^h \epsilon U_h$; and since $X_h = \emptyset$ (for $h\epsilon Q\backslash Q^*$), there also exists $\bar{u}^h \epsilon \bar{U}_h$ such that $\bar{u}^h b^h > 0$, $\forall\ h\epsilon Q\backslash Q^*$. But then, defining

$$u^h(\lambda) = \hat{u}^h + \lambda\bar{u}^h, \quad h\epsilon Q\backslash Q^* ,$$

for λ sufficiently large, $u^h(\lambda)\epsilon U_h$, $u^h(\lambda)b^h \geq z_o$, $\forall\ h\epsilon Q\backslash Q^*$.

Hence for all $h\epsilon Q$, there exist vectors u^h satisfying the constraints of (D) for $w = z_o$. To show that this is the maximal value of w, we note that since $\bar{x}$ is optimal for (P), there exists $h\epsilon Q$ such that

$$c\bar{x} = \min\{cx \mid x \in X_h\} .$$

But then by linear programming duality,

$$\begin{aligned} c\bar{x} &= \max\{u^h b^h \mid u^h \epsilon U_h\} \\ &= \max\{w \mid w - u^h b^h \leq 0,\ u^h \epsilon U_h\} \\ &\geq \max\{w \mid \bigwedge_{h\epsilon Q} (w - u^h b^h \leq 0,\ u^h \epsilon U_h)\} \end{aligned}$$

i.e., $w \leq z_o$, and hence the maximum value of w is $w_o = z_o$.

(ii) Assume that at least one of (P) and (D) is infeasible. If (P) is infeasible, $X_h = \emptyset$, $\forall\ h\epsilon Q$; hence for all $h\epsilon Q$, there exists $\bar{u}^h \epsilon \bar{U}_h$ such that $\bar{u}^h b^h > 0$.

If (D) is infeasible, we are done. Otherwise, for each $h\epsilon Q$ there exists $\hat{u}\epsilon U_h$. But then defining

$$u^h(\lambda) = \hat{u}^h + \lambda\bar{u}^h \quad , \quad h\epsilon Q,$$

$u^h(\lambda)\epsilon U_h$, $h\epsilon Q$, for all $\lambda > 0$, and since $\bar{u}^h b^h > 0$, $\forall\ h\epsilon Q$, w can be made arbitrarily large by increasing λ; i.e., (D) has no finite optimum.

Conversely, if (D) is infeasible, then either (P) is infeasible and we are done, or else, from the regularity condition, $Q^*\backslash Q^{**} \neq \emptyset$; and for $h\epsilon Q^*\backslash Q^{**}$ there exists $\hat{x} \in X_h$ and $\bar{x} \in \bar{X}_h$ such that $c\bar{x} < 0$. But then

$$x(\mu) = \hat{x} + \mu\bar{x}$$

is a feasible solution to (P) for any $\mu > 0$, and since $c\bar{x} < 0$, z can be made arbitrarily small by increasing μ; i.e., (P) has no finite optimum.

<u>Q.E.D.</u>

The above theorem asserts that either situation 1 or situation 2 holds for (P) and (D) <u>if</u> the regularity condition is satisfied. The following Corollary shows that the condition is not only sufficient but also necessary.

<u>Corollary 4.1.1</u>. If the regularity condition does not hold, then if (P) is feasible and (D) is infeasible, (P) has a finite minimum (i.e., there is a "duality gap").

<u>Proof</u>. Let (P) be feasible, (D) infeasible, and $Q^*\backslash Q^{**} = \emptyset$, i.e., for every $h \in Q^*$, let $U_h \neq \emptyset$. Then for each $h \in Q^*$, $\min\{cx \mid x \in X_h\}$ is finite, hence (P) has a finite minimum. <u>Q.E.D</u>.

<u>Remark</u>. The theorem remains true if some of the variables of (P) [of (D)] are unconstrained, and the corresponding constraints of (D) [of (P)] are equalities.

The regularity condition can be expected to hold in all but some rather peculiar situations. In linear programming duality, the case when both the primal and the dual problem is infeasible only occurs for problems whose coefficient matrix A has the rather special property that there exists $x \neq 0$, $u \neq 0$, satisfying the homogeneous system

$$Ax \geq 0 \quad , \quad x \geq 0$$

$$uA \leq 0 \quad , \quad u \geq 0$$

In this context, our regularity condition requires that, if the primal problem is feasible and the dual is infeasible, then at least one of the matrices A^h whose associated sets U_h are infeasible, should not have the above mentioned special property.

Though most problems satisfy this requirement, nevertheless there <u>are</u> situations when the regularity condition breaks down, as illustrated by the following example.

Consider the disjunctive program

$$\min -x_1-2x_2$$

$$\text{(P)}\qquad \begin{Bmatrix} -x_1+x_2 \geq 0 \\ -x_1-x_2 \geq -2 \\ x_1,x_2 \geq 0 \end{Bmatrix} \vee \begin{Bmatrix} -x_1+x_2 \geq 0 \\ x_1-x_2 \geq 1 \\ x_1,x_2 \geq 0 \end{Bmatrix}$$

and its dual

$$\text{(D)}\qquad \begin{array}{llll} \max w & & & \\ w & +2u_2^1 & & \leq 0 \\ & -u_1^1-u_2^1 & & \leq -1 \\ & u_1^1-u_2^1 & & \leq -2 \\ w & & -u_2^2 & \leq 0 \\ & & -u_1^2+u_2^2 & \leq -1 \\ & & u_1^2-u_2^2 & \leq -2 \\ & u_i^k \geq 0,\ i=1,2;\ k=1,2 & & \end{array}$$

The primal problem (P) has an optimal solution $\bar{x} = (0,2)$, with $c\bar{x} = -4$; whereas the dual problem (D) is infeasible. This is due to the fact that $Q\backslash Q^{**} = \{2\}$ and $X_2 = \emptyset$, $U_2 = \emptyset$, i.e., the regularity condition is violated. Here

$$X_2 = \left\{ x \in R_+^2 \;\middle|\; \begin{array}{l} -x_1+x_2 \geq 0 \\ x_1-x_2 \geq 1 \end{array} \right\},\quad U_2 = \left\{ u \in R_+^2 \;\middle|\; \begin{array}{l} -u_1^2+u_2^2 \leq -1 \\ u_1^2-u_2^2 \leq -2 \end{array} \right\}.$$

5. The Convex Hull of a Disjunctive Set

Having described the family of all valid inequalities, one is of course interested in identifying the strongest ones among the latter, i.e., the facets of the convex hull of feasible points of a disjunctive program.

If we denote the feasible set of a DP by

$$F = \left\{ x \in R^n \;\middle|\; \bigvee_{h \in Q} \begin{pmatrix} A^h x \geq a_0^h \\ x \geq 0 \end{pmatrix} \right\} ,$$

then for a given scalar α_0, the family of inequalities $\alpha x \geq \alpha_0$ satisifed by all $x \in F$, i.e., the family of valid inequalities for the DP, is obviously isomorphic to the family of vectors $\alpha \in F^{\#}_{(\alpha_0)}$, where

$$F^{\#}_{(\alpha_0)} = \{y \in R^n \mid yx \geq \alpha_0, \forall x \in F\} ,$$

in the sense that $\alpha x \geq \alpha_0$ is a valid inequality if and only if $\alpha \in F^{\#}_{(\alpha_0)}$.

In view of its relationship with ordinary polar sets, we call $F^{\#}_{(\alpha_0)}$ the <u>reverse polar</u> of F. Indeed, the ordinary polar set of F is

$$F^{o} = \{y \in R^n \mid yx \leq 1, \forall x \in F\} ,$$

and if we denote by $F^{o}_{(\alpha_0)}$ the polar of F scaled by α_0, (i.e., the set obtained by replacing 1 with α_0 in F), then $F^{\#}_{(\alpha_0)} = -F^{o}_{(-\alpha_0)}$.

The size (as opposed to the sign) of α_0 is of no interest here. Therefore we will distinguish only between the 3 cases $\alpha_0 > 0$ (or $\alpha_0 = 1$), $\alpha_0 = 0$ and $\alpha_0 < 0$ (or $\alpha_0 = -1$). (When the sign of α_0 makes no difference or is clear from the context, we will simply write $F^{\#}$.) For $\alpha_0 \leq 0$, as mentioned above, $F^{\#}_{(\alpha_0)}$ is (the negative of) an ordinary polar set, whose properties are described in the literature. The most interesting case for us, however, is $\alpha_0 > 1$, since this is the only case when the inequality $\alpha x \geq \alpha_0$ cuts off the point $x = 0$. This is why we need the concept of reverse polars.

For an arbitrary set $S \subseteq R^n$ we will denote by cl S, conv S, cone S, int S and dim S, the closure, the convex hull, the conical hull, the interior and the *dimension* of S, respectively. For a polyhedral set $S \subseteq R^n$ we will denote by vert S and dir S the set of vertices (extreme points) and the set of extreme direction vectors of S, respectively. For definitions and background material on these and related concepts (including ordinary polar sets), the reader is referred to [37] or [38] (see also [28]).

In [10] we showed that while some of the basic properties of polar sets carry over to reverse polars, others can only be recovered in a modified form. In the first category we mention (a) $(\lambda S)^{\#} = \frac{1}{\lambda} S^{\#}$; (b) $S \subseteq T \Rightarrow S^{\#} \supseteq T^{\#}$; (c) $(S \cup T)^{\#} = S^{\#} \cap T^{\#}$, properties which follow from the definitions. In the second one we state a few theorems, which are from [10] (see also [1]).

<u>Theorem 5.1</u>. (i) If $\alpha_0 \leq 0$, then $S^{\#} \neq \emptyset$ and

$$0 \in \text{int cl conv } S \iff S^{\#} \text{ is bounded}$$

(ii) If $\alpha_0 > 0$, then

$$0 \in \text{cl conv } S \iff S^{\#} = \emptyset \iff S^{\#} \text{ is bounded.}$$

<u>Proof</u>. (i) follows from the corresponding property of the ordinary polar S^o of S and the fact that $S^{\#}_{(\alpha_0)} = -S^o_{(\alpha_0)}$.

(ii) For $\alpha_0 > 0$, if $S^{\#} \neq \emptyset$ there exists $y \in R^n$ such that $xy \geq \alpha_0$, $\forall x \in \text{cl conv } S$. But $0 \cdot y < \alpha_0$, hence $0 \notin \text{cl conv } S$. Thus, if $0 \in \text{cl conv } S$, then $S^{\#} = \emptyset$; and hence $S^{\#}$ is bounded. Conversely, if $0 \notin \text{cl conv } S$, there exists a hyperplane $ax = \alpha_0$ separating 0 from cl conv S, i.e., such that ($\alpha_0 > 0$ and) $ax \geq \alpha_0$, $\forall x \in \text{cl conv } S$; which implies $a \in S^{\#}$, i.e., $S^{\#} \neq \emptyset$. It also implies $\lambda a \in S^{\#}, \forall \lambda > 1$, i.e., $S^{\#}$ is unbounded. ||

From this point on, we restrict our attention to sets S whose convex hull is polyhedral and pointed. For the disjunctive set F, this condition is satisfied if Q is finite. Most of the results carry over to the general case, but proofs are simpler with the above assumptions.

Theorem 5.2. If cl conv S is polyhedral, so is $S^{\#}$.

Proof. Let $u_1,\dots,u_p$ and $v_1,\dots,v_q$ be the vertices and extreme direction vectors, respectively, of cl conv S. Then for every $y \in S$ there exist scalars $\lambda_i \geq 0$, $i = 1,\dots,p$, $\mu_j \geq 0$, $j = 1,\dots,q$, with $\sum_{i=1}^{p} \lambda_i = 1$, such that

$$y = \sum_{i=1}^{p} u_i\lambda_i + \sum_{j=1}^{q} v_j\mu_j,$$

and it can easily be seen that for arbitrary x, $xy \geq \alpha_0$, $\forall\, y \in S$, if and only if $xu_i \geq \alpha_0$, $i = 1,\dots,p$, and $xv_j \geq 0$, $j = 1,\dots,q$. Thus

$$S^{\#} = \left\{ x \in R^n \middle| \begin{array}{ll} xu_i \geq \alpha_0 & , \quad i = 1,\dots,p \\ xv_j \geq 0 & , \quad j = 1,\dots,q \end{array} \right\}$$

i.e., $S^{\#}$ is polyhedral.||

The next result describes the crucial involutory property of polar and reverse polar sets.

Theorem 5.3. Assume $S^{\#} \neq \emptyset$. Then

$$S^{\#\#} = \begin{cases} \text{cl conv } S + \text{cl cone } S & \text{if } \alpha_0 > 0 \\ \text{cl cone } S & \text{if } \alpha_0 = 0 \\ \text{cl conv } (S \cup \{0\}) & \text{if } \alpha_0 < 0. \end{cases}$$

Proof. $S^{\#\#} = \{x \in R^n \mid xy \geq \alpha_0, \forall y \in S^{\#}\}$

$$= \left\{ x \in R^n \middle| \begin{array}{l} xy \geq \alpha_0 \text{ for all y s.t.} \\ u_i y \geq \alpha_0,\ i = 1,\dots,p \\ v_i y \geq 0,\ i = 1,\dots,q \end{array} \right\} .$$

But $xy \geq \alpha_0$ is a consequence of the system $u_i y \geq \alpha_0$, $i = 1,\dots,p$ and $v_i y \geq 0$, $i = 1,\dots,q$ (consistent, since $S^{\#} \neq \emptyset$), if and only if there exists a set of $\theta_i \geq 0$, $i = 1,\dots,p$, $\sigma_i \geq 0$, $i = 1,\dots,q$, such that

$$(5.1) \qquad x = \sum_{i=1}^{p} \theta_i u_i + \sum_{i=1}^{q} \sigma_i v_i ,$$

with

$$\sum_{i=1}^{p} \theta_i \alpha_0 \geq \alpha_0 .$$

Since $S^{\#}$ is polyhedral, so is $S^{\#\#}$. Thus $S^{\#\#}$ is the closed set of points $x \in R^n$ of the form (5.1) with $\theta_i \geq 0$, $i = 1,\dots,p$ $\sigma_i \geq 0$, $i = 1,\dots,q$, and

$$\sum_{i=1}^{p} \theta_i = \begin{cases} \geq 1 & \text{if } \alpha_0 > 0 \\ \geq 0 & \text{if } \alpha_0 = 0 \\ \leq 1 & \text{if } \alpha_0 < 0. \end{cases}$$

But these are precisely the expressions for the three sets claimed in the theorem to be equal to $S^{\#\#}$ in the respective cases. ||

<u>Corollary 5.3.1</u>. cl conv $S = S^{\#\#}_{(1)} \cap S^{\#\#}_{(-1)}$.

<u>Proof</u>. Follows immediately from the proof of Theorem 5.3, where $x \in S^{\#\#}_{(1)} \cap S^{\#\#}_{(-1)}$ corresponds to $\sum_{i=1}^{p} \theta_i = 1$.||

<u>Example 5.1.</u> Consider the disjunctive set

$$F = \left\{ x \in R^2 \;\middle|\; \begin{array}{l} -x_1 - x_2 \geq -2 \\ x_1 \qquad \geq 0 \\ \qquad x_2 \geq 0 \\ x_1 \geq 1 \vee x_2 \geq 1 \end{array} \right\}$$

illustrated in Fig. 5.1(a). Its reverse polars for $\alpha_0 = 1$ and $\alpha_0 = -1$ are the sets

$$F^{\#}_{(1)} = \left\{ y \in R^2 \;\middle|\; \begin{array}{rl} 2y_1 & \geq 1 \\ 2y_2 & \geq 1 \\ y_1 & \geq 1 \\ y_2 & \geq 1 \end{array} \right\} = \left\{ y \in R^2 \;\middle|\; \begin{array}{rl} y_1 & \geq 1 \\ y_2 & \geq 1 \end{array} \right\}$$

and

$$F^{\#}_{(-1)} = \left\{ y \in R^2 \;\middle|\; \begin{array}{rl} 2y_1 & \geq -1 \\ 2y_2 & \geq -1 \\ y_1 & \geq -1 \\ y_2 & \geq -1 \end{array} \right\} = \left\{ y \in R^2 \;\middle|\; \begin{array}{rl} 2y_1 & \geq -1 \\ 2y_2 & \geq -1 \end{array} \right\}$$

shown in Fig. 5.1(b) and (c).

Finally, the sets $F^{\#\#}$ corresponding to $\alpha_0 = 1$ and $\alpha_0 = -1$ (shown in Fig. 5.2(a), (b)) are

$$F^{\#\#}_{(1)} = \left\{ x \in R^2 \;\middle|\; \begin{array}{rl} x_1 + x_2 & \geq 1 \\ x_1 & \geq 0 \\ x_2 & \geq 0 \end{array} \right\}, \quad F^{\#\#}_{(-1)} = \left\{ x \in R^2 \;\middle|\; \begin{array}{rl} -x_1 - x_2 & \geq -2 \\ x_1 & \geq 0 \\ x_2 & \geq 0 \end{array} \right\}$$

and their intersection (shown in Fig. 5.2(c)) is

$$F^{\#\#}_{(1)} \cap F^{\#\#}_{(-1)} = \text{cl conv } F = \left\{ x \in R^2 \;\middle|\; \begin{array}{rl} -x_1 - x_2 & \geq -2 \\ x_1 + x_2 & \geq 1 \\ x_1 & \geq 0 \\ x_2 & \geq 0 \end{array} \right\} .$$

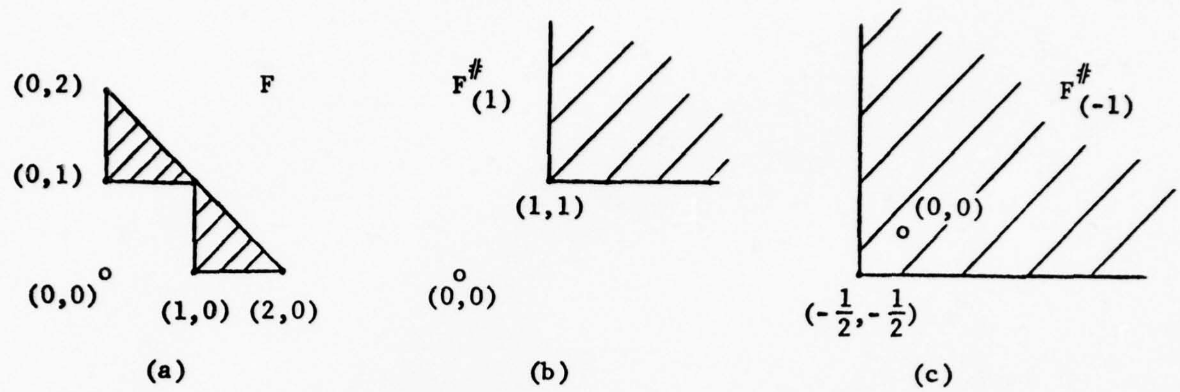


Fig. 5.1

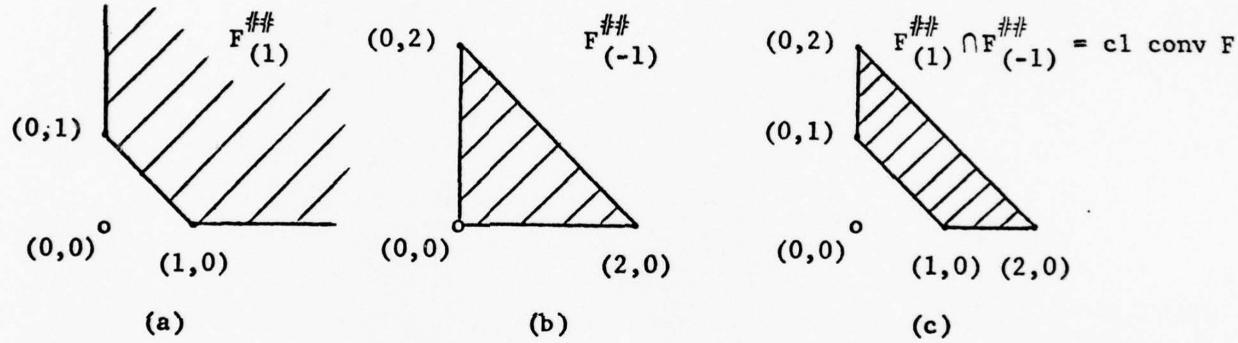


Fig. 5.2

The next theorem is needed to prove some other essential properties of reverse polars.

Theorem 5.4. $S^{\#\#\#} = S^{\#}$.

Proof. If $\alpha_0 \leq 0$, this follows from the corresponding property of ordinary polars. If $\alpha_0 > 0$ and $0 \in$ cl conv S, then $S^{\#} = \emptyset$, $S^{\#\#} = R^n$, and $S^{\#\#\#} = \emptyset = S^{\#}$. Finally, if $\alpha_0 > 0$ and $0 \notin$ cl conv S, then

$$\begin{aligned} S^{\#\#\#} &= \text{cl (conv } S + \text{cone } S)^{\#} \qquad \text{(from Theorem 5.3)} \\ &= \{y \in R^n \mid xy \geq \alpha_0,\ \forall x \in \text{cl (conv } S + \text{cone } S)\} \\ &= \{y \in R^n \mid xy \geq \alpha_0,\ \forall x \in S\} = S^{\#}. \ \| \end{aligned}$$

The above results can be used to characterize the facets of cl conv S. To simplify the exposition, we will assume that S is a full-dimensional pointed polyhedron, which implies that $S^{\#\#}$ is also full-dimensional. For the general case the reader is referred to [10].

We recall that an inequality $\pi x \geq \pi_0$ defines a facet of a convex set C if $\pi x \geq \pi_0$, $\forall x \in C$, and $\pi x = \pi_0$ for exactly d affinely independent points x of C, where d = dim C. The facet of C defined by $\pi x \geq \pi_0$ is then $\{x \in C \mid \pi x = \pi_0\}$; but as customary in the literature, for the sake of brevity we call $\pi x \geq \pi_0$ itself a facet.

We proceed in two steps, the first of which concerns the facets of $S^{\#\#}$.

Theorem 5.5. Let dim S = n, and $\alpha_0 \neq 0$. Then $\alpha x \geq \alpha_0$ is a facet of $S^{\#\#}$ if and only if $\alpha \neq 0$ is a vertex of $S^{\#}$.

Proof. From Theorem 5.4, $\alpha \in R^n$ is a vertex of $S^{\#}$ if and only if

$$\alpha \in S^{\#} = \left\{ y \in R^n \;\middle|\; \begin{array}{l} uy \geq \alpha_0, \ \forall u \in \text{vert } S^{\#\#} \\ vy \geq 0, \ \forall v \in \text{dir } S^{\#\#} \end{array} \right\}$$

and α satisfies with equality a subset of rank n of the system defining $S^{\#}$. Further, $\alpha \neq 0$ if and only if this subset of inequalities is not homogeneous (i.e., at least one right hand side coefficient is $\alpha_0 \neq 0$).

On the other hand, $\alpha x \geq \alpha_0$ is a facet of $S^{\#\#}$ if and only if (i) $\alpha x \geq \alpha_0$, $\forall x \in S^{\#\#}$, i.e., $\alpha \in S^{\#}$; and (ii) $\alpha x = \alpha_0$ for exactly n affinely independent points of $S^{\#\#}$. But (ii) holds if and only if $\alpha u = \alpha_0$ for r vertices u of $S^{\#\#}$, and $\alpha v = 0$ for s extreme direction vectors v of $S^{\#\#}$, with $r \geq 1$ (since $\alpha \neq 0$) and $r + s \geq n$, such that the system of these equations is of rank n.

Thus the two sets of conditions (for $\alpha x \geq \alpha_0$ to be a facet of $S^{\#\#}$ and for α to be a vertex of $S^{\#}$) are identical. ||

By arguments similar to the above proof, one shows that $\alpha x \geq 0$ is a facet of $S^{\#\#}$ if and only if α is an extreme direction vector of $S^{\#}$. Unlike for $\alpha_0 \neq 0$, the homogeneous inequality $\alpha x \geq 0$ is a facet of $S^{\#\#}_{(1)}$ if and only if it is also a facet of $S^{\#\#}_{(-1)}$ [of $S^{\#\#}_{(0)}$], due to the fact that every extreme direction vector of $S^{\#}_{(1)}$ is also an extreme direction vector of $S^{\#}_{(-1)}$ [of $S^{\#}_{(0)}$], and vice-versa.

<u>Theorem 5.6</u>. Let dim S = n, and $\alpha_0 \neq 0$. Then $\alpha x \geq \alpha_0$ is a facet of cl conv S if and only if it is a facet of $S^{\#\#}_{(\alpha_0)}$.

<u>Proof</u>. (i) If $\alpha_0 > 0$, the halfspace $\alpha x \geq \alpha_0$ contains cl conv S if and only if it contains cl(conv S + cone S). If $\alpha_0 < 0$, the halfspace $\alpha x \geq \alpha_0$ contains cl conv S if and only if it contains cl conv $(S \cup \{0\})$. From Theorem 4.3, in both cases $\alpha x \geq \alpha_0$ is a supporting halfspace for cl conv S if and only if it is a supporting halfspace for $S^{\#\#}_{(\alpha_0)}$.

(ii) Next we show that

$$\{x \in \text{cl conv } S \mid \alpha x = \alpha_0\} = \{x \in S^{\#\#}_{(\alpha_0)} \mid \alpha x = \alpha_0\} .$$

The relation $\subseteq$ follows from cl conv $S \subseteq S^{\#\#}_{(\alpha_0)}$ (Theorem 5.3). To show the converse, assume it to be false, and let $x \in S^{\#\#}_{(\alpha_0)} \setminus$cl conv S satisfy $\alpha x = \alpha_0$. From Theorem 5.3, $x = \lambda u$ for some $u \in$ cl conv S, and $\lambda > 1$ if $\alpha_0 > 0$, $0 < \lambda < 1$ if $\alpha_0 < 0$. In each case, $\alpha x = \alpha_0$ implies $\alpha u = (1/\lambda)\alpha x < \alpha_0$ for some $u \in$ cl conv $S \subseteq S^{\#\#}_{(\alpha_0)}$, contrary to the assumption that $\alpha x \geq \alpha_0, \forall\, x \in S^{\#\#}_{(\alpha_0)}$. ||

By an argument similar to the above proof one can show that if $\alpha x \geq 0$ is a facet of cl conv S, then it is a facet of $S^{\#\#}_{(\alpha_0)}$ for both $\alpha_0 = 1$ and $\alpha_0 = -1$. The converse, however, is not true, i.e., $\alpha x \geq 0$ can be a facet of both $S^{\#\#}_{(1)}$ and $S^{\#\#}_{(-1)}$, without being a facet of $S^{\#\#}_{(1)} \cap S^{\#\#}_{(-1)}$.

We are now ready to characterize the facets of the convex hull of the disjunctive set

$$F = \left\{ x \in R^n \;\middle|\; \bigvee_{h \in Q} \begin{pmatrix} A^h x \geq a_0^h \\ x \geq 0 \end{pmatrix} \right\},$$

where Q is assumed to be finite, and F to be full-dimensional (for the general case see [10]).

<u>Theorem 5.7</u>. $\alpha x \geq \alpha_0$, with $\alpha_0 \neq 0$, is a facet of cl conv F *if and only if* $\alpha \neq 0$ is a vertex of the polyhedron

$$F^{\#}_{(\alpha_0)} = \left\{ y \in R^n \;\middle|\; \begin{array}{l} y \geq \theta^h A^h,\ h \in Q^* \\ \text{for some } \theta^h \geq 0,\ h \in Q^* \\ \text{satisfying } \theta^h a_0^h \geq \alpha_0 \end{array} \right\},$$

where Q^* *is the set of those* $h \in Q$ for which the system $A^h x \geq a_0^h$, $x \geq 0$, is consistent.

<u>Proof</u>. From Theorem 3.1, the set $F^{\#}_{(\alpha_0)} = \{y \in R^n \mid xy \geq \alpha_0,\ \forall x \in F\}$ is of the form claimed above. The rest is a direct application of Theorems 5.5 and 5.6. ||

As for the case $\alpha_0 = 0$, from the above comments it follows that if $\alpha x \geq 0$ is a facet of cl conv F, then $\alpha \neq 0$ is an extreme direction vector of $F^{\#}_{(\alpha_0)}$ for all α_0. The converse is not true, but if $\alpha \neq 0$ is an extreme direction vector of $F^{\#}_{(\alpha_0)}$ for some α_0 (hence for all α_0), then $\alpha x \geq 0$ is either a facet of cl conv F, or the intersection of two facets, $\alpha^1 x \geq \alpha_0^1$ and $\alpha^2 x \geq \alpha_0^2$, with $\alpha^1 + \alpha^2 = \alpha$ and $\alpha_0^1 = -\alpha_0^2 \neq 0$ (see [10] for the details).

Since the facets of cl conv F are vertices (in the nonhomogeneous case) or extreme direction vectors (in the homogeneous case) of the convex polyhedron $F^{\#}$, they can be found by maximizing or minimizing some properly chosen linear function on $F^{\#}$, i.e., by solving a linear program of the form

$$P_1^*(g,\alpha_0) \qquad \begin{array}{rcll} \min gy & & & \\ y - \theta^h A^h & \geq & 0 & \\ \theta^h a_0^h & \geq & \alpha_0 & h \in Q^* \\ \theta^h & \geq & 0 & \end{array}$$

or its dual

$$P_2^*(g,\alpha_0) \qquad \begin{array}{rcll} \max \sum_{h \in Q^*} \alpha_0 \xi_0^h & & & \\ \sum_{h \in Q^*} \xi^h & = & g & \\ a_0^h \xi_0^h - A^h \xi^h & \leq & 0 & h \in Q^* \\ \xi_0^h \geq 0, \ \xi^h \geq 0 & & & \end{array}$$

From Theorem 5.1, if $\alpha_0 \leq 0$ then $F^{\#}_{(\alpha_0)} \neq \emptyset$, i.e., $P_1^*(g,\alpha_0)$ is always feasible; whereas if $\alpha_0 > 0$, then $P_1^*(g,\alpha_0)$ is feasible if and only if $0 \notin$ cl conv F. This latter condition expresses the obvious fact that an inequality which cuts off the origin can only be derived from a disjunction which itself cuts off the origin.

Two problems arise in connection with the use of the above linear programs to generate facets of cl conv F. The first one is that sometimes only Q is known, but Q^* is not. This can be taken care of by working with Q rather than Q^*. Let $P_k(g,\alpha_0)$ denote the problem obtained by replacing Q^* with Q in $P_k^*(g,\alpha_0)$, k = 1,2. It was shown in [10], that if $P_2(g,\alpha_0)$ has an optimal solution $\bar{\xi}$ such that

$$(\bar{\xi}_0^h = 0, \ \bar{\xi}^h \neq 0) \Rightarrow h \in Q^* \ ,$$

Name: ______

Finance Quiz
Supplementary Page.

3. GIVEN THE OPTION OF PURCHASING ANY ONE OF TWO STOCKS A AND B WITH THE FOLLOWING RISK-RETURN CHARACTERISTICS, JOHN SAYS HE WILL BUY STOCK A. FROM THIS DATA I CAN CONCLUDE THAT:

	EXPECTED RETURN	RISK
STOCK A	5%	8%
STOCK B	6%	9%

A. JOHN IS A RISK LOVER
B. JOHN IS RISK ADVERSE.
C. JOHN IS NEUTRAL TOWARDS RISK
D. JOHN IS IRRATIONAL
E. NONE OF THE ABOVE.

4. NOW SUPPOSE THAT JOHN HAS THE CHOICE OF ANY ONE OF THREE STOCKS A,B AND C. A AND B HAVE THE SAME RISK-RETURN CHARACTERISTICS AS IN QUESTION 3. STOCK C HAS EXPECTED RETURN OF 5% AND RISK OF 9%. WHICH STOCK WILL JOHN CHOOSE?

A. STOCK A
B. STOCK B
C. STOCK C
D. UNDETERMINED WITH GIVEN INFORMATION
E. HE WOULD BE INDIFFERENT BETWEEN STOCKS B AND C

then every optimal solution of $P_1(g,\alpha_0)$ is an optimal solution of $P_1^*(g,\alpha_0)$. Thus, one can start by solving $P_2(g,\alpha_0)$. If the above condition is violated for some $h \in Q\setminus Q^*$, then h can be removed from Q and $P_2(g,\alpha_0)$ solved for Q redefined in this way. When necessary, this procedure can be repeated.

The second problem is that, since the facets of cl conv F of primary interest are the nonhomogeneous ones (in particular those with $\alpha_0 > 0$, since they cut off the origin), one would like to identify the class of vectors g for which $P_1^*(g,\alpha_0)$ has a finite minimum. It was shown in [10], that $P_1^*(g,\alpha_0)$ has a finite minimum if and only if $\lambda g \in$ cl conv F for some $\lambda > 0$; and that, should g satisfy this condition, $\alpha x \geq \alpha_0$ is a facet of cl conv F (where F is again assumed full-dimensional) if and only if $\alpha = \bar{y}$ for every optimal solution $(\bar{y},\bar{\theta})$ to $P_1^*(g,\alpha_0)$.

As a result of these characterizations, facets of the convex hull of the feasible set F can be computed by solving the linear program $P_1(g,\alpha_0)$ or its dual. If the disjunction defining F has many terms, like in the case where F comes from the disjunctive programming formulation of a 0-1 program with a sizeable number of 0-1 conditions, $P_1(g,\alpha_0)$ is too large to be worth solving. If, however, F is made to correspond to a relaxation of the original zero-one program, involving zero-one conditions for only a few well chosen variables, then $P_1(g,\alpha_0)$ or its dual is practically solvable and provides the strongest possible cuts obtainable from those particular zero-one conditions.

On the other hand, since the constraint set of $P_2(g,\alpha_0)$ consists of $|Q|$ more or less loosely connected subsystems, one is tempted to try to approximate an optimal solution to $P_2(g,\alpha_0)$ — and thereby to $P_1(g,\alpha_0)$ — by solving the subsystems independently. Early computational experience indicates that these approximations are quite good.

We now give a numerical example for a facet calculation.

<u>Example 5.2</u>. Find all those facets of cl conv F which cut off the origin (i.e., all facets of the form $\alpha x \geq 1$), where $F \subset R^2$ is the disjunctive set

$$F = F_1 \vee F_2 \vee F_3 \vee F_4 ,$$

with

$$F_1 = \{x \mid -x_1 + 2x_2 \geq 6,\ 0 \leq x_1 \leq 1,\ x_2 \geq 0\}$$

$$F_2 = \{x \mid 4x_1 + 2x_2 \geq 11,\ 1 \leq x_1 \leq 2.5,\ x_2 \geq 0\}$$

$$F_3 = \{x \mid -x_1 + x_2 \geq -2,\ 2.5 \leq x_1 \leq 4,\ x_2 \geq 0\}$$

$$F_4 = \{x \mid x_1 + x_2 \geq 6,\ 4 \leq x_1 \leq 6,\ x_2 \geq 0\}$$

(see Fig. 5.3).

After removing some redundant constraints, F can be restated as the set of those $x \in R^2_+$ satisfying

$$\{-x_1 + 2x_2 \geq 6\} \ \vee \ \left\{\begin{matrix} 4x_1 + 2x_2 \geq 11 \\ -x_1 + x_2 \geq -2 \end{matrix}\right\} \ \vee \ \{x_1 + x_2 \geq 6\}$$

and the corresponding problem $P_1(g,1)$ is

$$\begin{array}{lllllllll}
\min g_1y_1 + g_2y_2 & & & & & & & & \\
y_1 & & + \theta_1^1 & & & & & \geq & 0 \\
 & y_2 & - 2\theta_1^1 & & & & & \geq & 0 \\
y_1 & & & - 4\theta_1^2 & + \theta_2^2 & & & \geq & 0 \\
 & y_2 & & - 2\theta_1^2 & - \theta_2^2 & & & \geq & 0 \\
y_1 & & & & & & - \theta_1^3 & \geq & 0 \\
 & y_2 & & & & & - \theta_1^3 & \geq & 0 \\
 & & 6\theta_1^1 & & & & & \geq & 1 \\
 & & & 11\theta_1^2 & - 2\theta_2^2 & & & \geq & 1 \\
 & & & & & & 6\theta_1^3 & \geq & 1 \\
 & & \multicolumn{5}{l}{\theta_1^1,\ \theta_1^2,\ \theta_2^2,\ \theta_1^3 \geq 0 .} & &
\end{array}$$

Solving this linear program for $g = (1,1)$, yields the optimal points $(y;\theta) = \left(\frac{1}{3},\frac{1}{3};\frac{1}{6},\frac{1}{9},\frac{1}{9},\frac{1}{6}\right)$, and $(y;\theta) = \left(\frac{1}{3},\frac{1}{3};\frac{1}{6},\frac{1}{9},\frac{1}{9},\frac{1}{3}\right)$, which have the same y-component: $\left(\frac{1}{3},\frac{1}{3}\right)$. These points are optimal (and the associated y is unique) for all $g > 0$ such that $g_1 < 5g_2$. For $g_1 = 5g_2$, in addition to the above points, which are still optimal, the points $(y;\theta) = \left(\frac{1}{6},\frac{7}{6};\frac{1}{6},\frac{2}{9},\frac{13}{18},\frac{1}{6}\right)$ and $(y;\theta) = \left(\frac{1}{6},\frac{7}{6};\frac{7}{12},\frac{2}{9},\frac{13}{18},\frac{1}{6}\right)$, which again have the same y-component $y = \left(\frac{1}{6},\frac{7}{6}\right)$, also become optimal; and they are the only optimal solutions for all $g > 0$ such that $g_1 > 5g_2$.

We have thus found that the convex hull of F has two facets which cut off the origin, corresponding to the two vertices $y^1 = \left(\frac{1}{3},\frac{1}{3}\right)$ and $y^2 = \left(\frac{1}{6},\frac{7}{6}\right)$ of $F^{\#}(1)$:

$$\frac{1}{3}x_1 + \frac{1}{3}x_2 \geq 1$$

$$\frac{1}{6}x_1 + \frac{7}{6}x_2 \geq 1$$

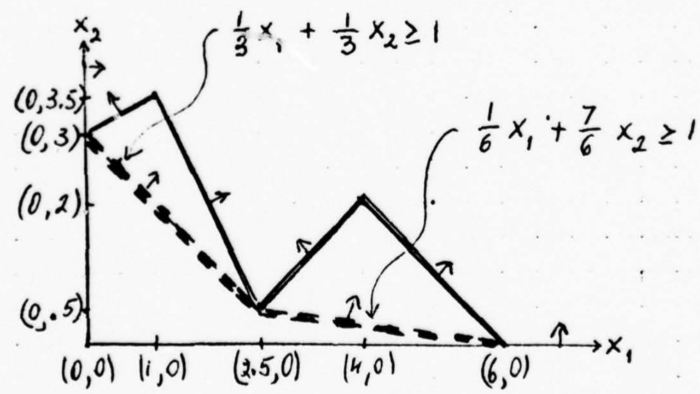


Fig. 5.3

6. Facial Disjunctive Programs

In this section we discuss the following problem [10]. Given a disjunctive program in the conjunctive normal form (2.2), is it possible to generate the convex hull of feasible points by imposing the disjunctions $j \in S$ one by one, at each step calculating a "partial" convex hull, i.e., the convex hull of the set defined by the inequalities generated earlier, plus one of the disjunctions?

For instance, in the case of an integer program, is it possible to generate the convex hull of feasible points by first producing all the facets of the convex hull of points satisfying the linear inequalities, plus the integrality condition on, say, x_1; then adding all these facet-inequalities to the constraint set and generating the facets of the convex hull of points satisfying this amended set of inequalities, plus the integrality condition on x_2; etc. The question has obvious practical importance, since calculating facets of the convex hull of points satisfying one disjunction is a considerably easier task, as shown in the previous section, than calculating facets of the convex hull of the full disjunctive set.

The answer to the above question is negative in general, but positive for a very important class of disjunctive programs, which we term facial. The class includes (pure or mixed) 0-1 programs.

The fact that in the general case the above procedure does not produce the convex hull of the feasible points can be illustrated on the following 2-variable problem.

Example 6.1. Given the set

$$F_0 = \{x \in R^2 \mid -2x_1 + 2x_2 \le 1,\ 2x_1 - 2x_2 \le 1,\ \ 0 \le x_1 \le 2,\ 0 \le x_2 \le 2\}$$

find $F = \text{cl conv}\ (F_0 \cap \{x \mid x_1, x_2 \text{ integer}\})$.

Denoting

$$F_1 = \text{cl conv}\,(F_0 \cap \{x \mid x_1 \text{ integer}\}), \quad F_2 = \text{cl conv}\,(F_1 \cap \{x \mid x_2 \text{ integer}\}),$$

the question is whether $F_2 = F$. As shown in Fig. 6.1, the answer is no, since

$$F_2 = \left\{ x \;\middle|\; \begin{array}{r} 2x_1 - x_2 \geq 0 \\ -2x_1 + 3x_2 \geq 0 \\ -2x_1 + x_2 \geq -2 \\ 2x_1 - 3x_2 \geq -2 \end{array} \right\}, \text{ while } F = \left\{ x \;\middle|\; \begin{array}{c} -x_1 + x_2 = 0 \\ 0 \leq x_1 \leq 2 \\ 0 \leq x_2 \leq 2 \end{array} \right\}.$$

If the order in which the integrality constraints are imposed is reversed, the outcome remains the same.

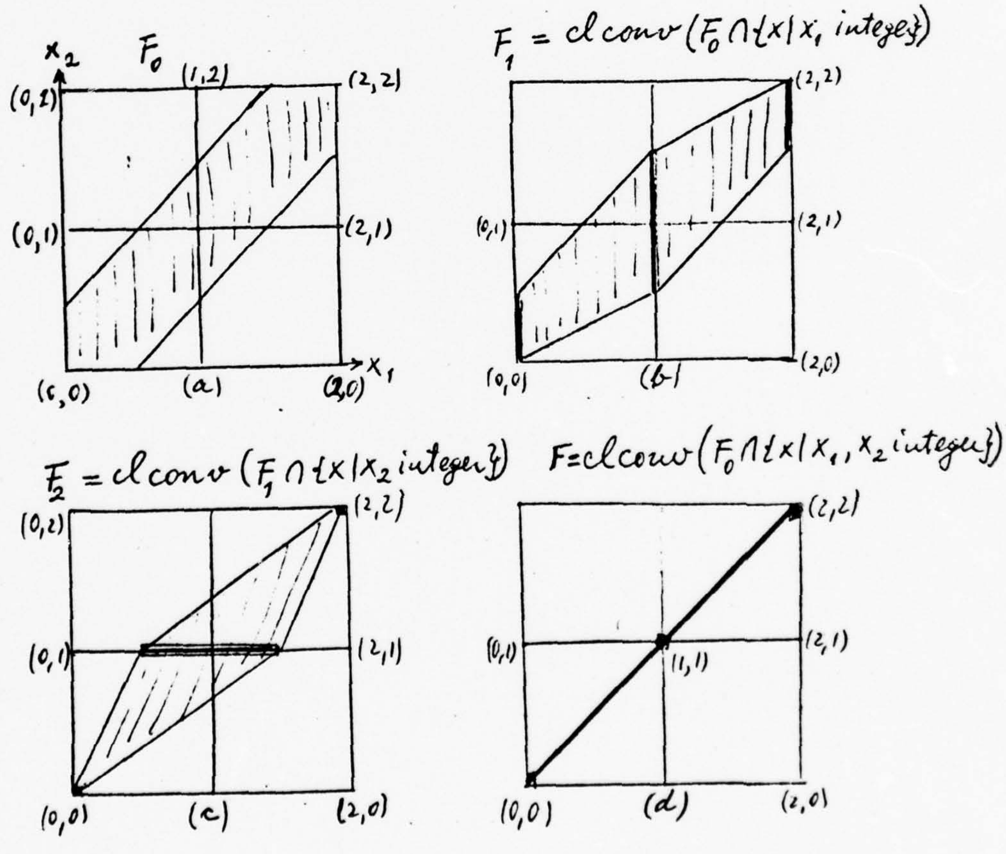


Fig. 6.1

Consider a disjunctive program stated in the conjunctive normal form (2.2), and denote

$$F_0 = \{x \in R^n \mid Ax \geq a_0, \; x \geq 0\} \; .$$

The disjunctive program (and its constraint set) is called facial if every inequality $d^i x \geq d_{i0}$ that appears in a disjunction of (2.2), defines a face of F_0; i.e., if for all $i \in Q_j$, $j \in S$, the set

$$F_0 \cap \{x \mid d^i x \geq d_{i0}\}$$

is a face of F_0. (A face of a polyhedron P is the intersection of P with some of its boundary planes.)

The class of disjunctive programs that have the facial property includes the most important cases of disjunctive programming, like the 0-1 programming (pure or mixed), nonconvex quadratic programming, separable programming, the linear complementarity problem, etc.; but not the general integer programming problem, as illustrated above. In all the above mentioned cases the inequalities $d^i x \geq d_{i0}$ of each disjunction actually define facets, i.e., $(d - 1)$-dimensional faces of F_0, where d is the dimension of F_0.

Another property that we need is the boundedness of F_0. Since this can always be achieved, if necessary, by regularizing F_0, its assumption does not represent a practical limitation.

Theorem 6.1. Let the constraint set of a DP be represented as

$$F = \{x \in F_0 \mid \bigvee_{i \in Q_j} (d^i x \geq d_{i0}), \; j \in S\}$$

where

$$F_0 = \{x \in R^n \mid Ax \geq a_0, \; x \geq 0\} \; .$$

For an arbitrary ordering of S, define recursively

$$F_j = \text{conv} \left[\bigvee_{i \in Q_j} (F_{j-1} \cap \{x \mid d^i x \geq d_{i0}\}) \right], \quad j = 1, \ldots, |S|.$$

If F is facial and F_0 is bounded, then

$$F_{|S|} = \text{conv } F .$$

The proof of this theorem [10] uses the following auxiliary result.

<u>Lemma 6.1.1.</u> Let $P_1, \ldots, P_r$ be a finite set of polytopes (bounded polyhedra), and $P = \bigcup_{h=1}^{r} P_h$.

Let $H^+ = \{x \in R^n \mid d^i x \leq d_{i0}\}$ be an arbitrary halfspace, and $H = \{x \in R^n \mid d^i x = d_{i0}\}$ its defining hyperplane.

If $P \subseteq H^+$, then

$$H \cap \text{conv } P = \text{conv } (H \cap P) .$$

<u>Proof</u>. Let $H \cap \text{conv } P \neq \emptyset$ (otherwise the Lemma holds trivially). Clearly, $(H \cap P) \subseteq (H \cap \text{conv } P)$, and therefore

$$\text{conv } (H \cap P) \subseteq \text{conv } (H \cap \text{conv } P) = H \cap \text{conv } P .$$

To prove $\supseteq$, let $u_1, \ldots, u_p$ be the vertices of all the polytopes P_h, $h = 1, \ldots, r$. Obviously, p is finite, conv P is closed, and vert conv $P \subseteq (\bigcup_{h=1}^{r} \text{vert } P_h)$. Then

$$x \in H \cap \text{conv } P \Rightarrow d^i x = d_{i0} \text{ and } x = \sum_{k=1}^{p} \lambda_k u_k ,$$

with

$$\sum_{k=1}^{p} \lambda_k = 1, \quad \lambda_k \geq 0, \quad k = 1, \ldots, p .$$

Further, $P \subseteq H^+$ implies $d^i u_k \leq d_{i0}$, $k = 1, \ldots, p$. We claim that in the above expression for x, if $\lambda_k > 0$ then $d^i u_k = d_{i0}$. To show this, we assume there exist $\lambda_k > 0$ such that $d^i u_k < d_{i0}$. Then

$$d^i x = d^i(\sum_{k=1}^{p} \lambda_k u_k) < d_{i0}(\sum_{k=1}^{p} \lambda_k)$$

$$= d_{i0} ,$$

a contradiction. Hence x is the convex combination of points $u^k \in H \cap P$, or $x \in \text{conv}\,(H \cap P)$. ||

Another relation to be used in the proof of the theorem, is the fact that for arbitrary $S_1, S_2 \subseteq R^n$,

(6.1) $$\text{conv}\,(\text{conv}\,S_1 \cup \text{conv}\,S_2) = \text{conv}\,(S_1 \cup S_2) .$$

Proof of Theorem 6.1. For $j = 1$, the statement is true from the definitions and the obvious relation

$$\bigvee_{i\in Q_1} (F_0 \cap \{x \mid d^i x \geq d_{i0}\}) = \{x \in F_0 \mid \bigvee_{i\in Q_1} (d^i x \geq d_{i0})\} .$$

To prove the theorem by induction on j, suppose the statement is true for $j = 1,\ldots,k$. Then

$$F_{k+1} = \text{conv}\,[\bigvee_{i\in Q_{k+1}} (F_k \cap \{x \mid d^i x \geq d_{i0}\})] \qquad \text{(by definition)}$$

$$= \text{conv}\,[\bigvee_{i\in Q_{k+1}} (\{x \mid d^i x \geq d_{i0}\} \cap \text{conv}\,\{x \in F_0 \mid \bigvee_{i\in Q_j} (d^i x \geq d_{i0}),\ j = 1,\ldots,k\})$$

(from the assumption)

$$= \text{conv}\,[\bigvee_{i\in Q_{k+1}} \text{conv}\,(\{x \mid d^i x \geq d_{i0}\} \cap \{x \in F_0 \mid \bigvee_{i\in Q_j} (d^i x \geq d_{i0}),\ j = 1,\ldots,k\})]$$

(from Lemma 6.1.1)

$$= \text{conv}\,[(\bigvee_{i\in Q_{k+1}} \{x \mid d^i x \geq d_{i0}\}) \cap \{x \in F_0 \mid \bigvee_{i\in Q_j} (d^i x \geq d_{i0}),\ j = 1,\ldots,k\}]$$

(from (6.1))

$$= \text{conv}\,\{x \in F_0 \mid \bigvee_{i\in Q_j} (d^i x \geq d_{i0}),\ j = 1,\ldots,k+1\} ,$$

i.e., the statement is also true for $j = k + 1$. ||

Theorem 6.1 implies that for a bounded facial disjunctive program with feasible set F, the convex hull of F can be generated in $|S|$ stages, (where S is as in (2.2)), by generating at each stage a "partial" convex hull, namely the convex hull of a disjunctive program with only one disjunction.

In terms of a 0-1 program, for instance, the above result means that the problem

$$\min \{cx \mid Ax \geq b,\ 0 \leq x \leq e,\ x_j = 0 \text{ or } 1,\ j = 1,\ldots,n\} ,$$

where $e = (1,\ldots,1)$, is equivalent to (has the same convex hull of its feasible points, as)

$$\min \{cx \mid Ax \geq b,\ 0 \leq x \leq e,\ \alpha^i x \geq \alpha_{i0},\ i \in H_1,\ x_j = 0 \text{ or } 1,\ j = 2,\ldots,n\} \tag{6.2}$$

where $\alpha^i x \geq \alpha_{i0}$, $i \in H_1$, are the facets of

$$F_1 = \operatorname{conv} \{x \mid Ax \geq b,\ 0 \leq x \leq e,\ x_1 = 0 \text{ or } 1\} .$$

In other words, x_1 is guaranteed to be integer-valued in a solution of (6.2) although the condition $x_1 = 0$ or 1 is not present among the constraints of (6.2). A 0-1 program in n variables can thus be replaced by one in n-1 variables at the cost of introducing new linear inequalities. The inequalities themselves are not expensive to generate, since the disjunction that gives rise to them $(x_1 = 0 \vee x_1 = 1)$ has only two terms. The difficulty lies rather in the number of facets that one would have to generate, were one to use this approach for solving 0-1 programs. However, by using some information as to which inequalities (facets of a "partial" convex hull) are likely to be binding at the optimum, one might be able to make the above approach efficient by generating only a few facets of the "partial" convex hull at each iteration. This question requires further investigation. For additional results on facial disjunctive programs see [33], [34].

7. Disjunctive Programs with Explicit Integrality Constraints

The theory reviewed in the previous sections derives cutting planes from disjunctions. In this context, 0-1 or integrality conditions are viewed as disjunctions, and the disjunction to be used for deriving a cut usually applies to the basic variables.

In this section we discuss a principle for strengthening cutting planes derived from disjunctions in the case when, besides the disjunction which applies to the basic variables, there are also integrality constraints on some of the nonbasic variables. In [14] we first proved this principle for arbitrary cuts, by using subadditive functions, then applied it to cuts from disjunctions. Here we prove the principle directly for the latter case, without recourse to concepts outside the framework of disjunctive programming.

Let a DP be stated in the disjunctive normal form (2.1), and assume in addition that some components of x are integer-constrained. In order for the principle that we are going to discuss to be applicable, it is necessary that each $A^h x$, $h \in Q$, has a lower bound, say b_0^h. With these additional features, and denoting by J the index set for the components of x ($|J| = n$), the constraint set of the DP can be stated as

(7.1) $$A^i x \geq b_0^i, \quad i \in Q$$
$$x \geq 0 ,$$

(7.2) $$\bigvee_{i \in Q} (A^i x \geq a_0^i)$$

and

(7.3) $$x_j \text{ integer}, \quad j \in J_1 \subset J ,$$

where

(7.4) $$a_0^i \geq b_0^i, \quad i \in Q .$$

Let $Q = \{1,\ldots,q\}$, and let a_j^i stand for the j-th column of A^i, $j \in J$, $i \in Q$.

<u>Theorem 7.1</u>. [14] Define

(7.5) $$M = \{m \in R^q \mid \sum_{i \in Q} m_i \geq 0,\ m_i \text{ integer},\ i \in Q\} .$$

Then every $x \in R^n$ that satisfies (7.1), (7.2), (7.3), also satisfies the inequality

(7.6) $$\sum_{j \in J} \alpha_j x_j \geq \alpha_0 ,$$

where

(7.7) $$\alpha_j = \begin{cases} \inf\limits_{m \in M} \max\limits_{i \in Q} \theta^i [a_j^i + m_i(a_0^i - b_0^i)] , & j \in J_1 \\ \\ \max\limits_{i \in Q} \theta^i a_j^i & j \in J \setminus J_1 = J_2 \end{cases}$$

and

(7.8) $$\alpha_0 = \min_{i \in Q} \theta^i a_0^i .$$

To prove this theorem we will use the following auxiliary result.

<u>Lemma 7.1</u>. Let $m_j \in M$, $m_j = (m_{ij})$, $j \in J_1$. Then for every $x \in R^n$ satisfying (7.3), either

(7.9) $$\sum_{j \in J_1} m_{ij} x_j = 0 , \quad \forall i \in Q$$

or

(7.10) $$\bigvee_{i \in Q} \Big(\sum_{j \in J_1} m_{ij} x_j \geq 1 \Big)$$

holds.

Proof. If the statement is false, there exists $\bar{x}$ satisfying (7.3) and such that

$$\sum_{i\in Q} \sum_{j\in J_1} m_{ij}\bar{x}_j < 0 .$$

On the other hand, from $\bar{x} \geq 0$ and the definition of M,

$$\sum_{i\in Q} \sum_{j\in J_1} m_{ij}\bar{x}_j \geq 0 ,$$

a contradiction. ||

Proof of Theorem 7.1. We first show that every x which satisfies (7.1), (7.2) and (7.3), also satisfies

$$(7.2') \qquad \bigvee_{i\in Q}\left[\sum_{j\in J_1} [a_j^i + m_{ij}(a_0^i - b_0^i)]x_j + \sum_{j\in J_2} a_j^i x_j \geq a_0^i\right]$$

for any set of $m_j \in M$, $j \in J_1$. *To see this, write (6.2′) as*

$$(7.2'') \qquad \bigvee_{i\in Q} [\sum_{j\in J} a_j^i x_j + (a_0^i - b_0^i) \sum_{j\in J_1} m_{ij}x_j \geq a_0^i] .$$

From Lemma 7.1, either (7.9) or (7.10) holds for every x satisfying (7.3). If (7.9) holds, then (7.2″) is the same as (7.2) which holds by assumption. If (7.10) holds, there exists $k \in Q$ such that $\sum_{j\in J_1} m_{kj}x_j = 1 + \lambda$ for some $\lambda \geq 0$. But then the k-th term of (7.2″) becomes

$$\sum_{j\in J} a_j^k x_j \geq b_0^k - \lambda(a_0^k - b_0^k)$$

which is satisfied since $\lambda(a_0^k - b_0^k) \geq 0$ and x satisfies (7.1). This proves that every feasible x satisfies (7.2′).

Applying to (7.2′) Theorem 3.1 then produces the cut (7.6) with coefficients defined by (7.7), (7.8). Taking the infinum over M is justified by the fact that (7.6) is valid with α_0 as in (7.8), α_j as in (7.7) for $j \in J_2$, and

$$\alpha_j = \max_{i \in Q} \theta^i [a_j^i + m_{ij}(a_0^i - b_0^i)]$$

for $j \in J_1$, for arbitrary $m_j \in M$. ||

<u>Corollary 7.1.1</u>. [14]. Let the vectors σ^i, $i \in Q$, satisfy

$$\sigma^i(a_0^i - b_0^i) = 1, \quad \sigma^i a_0^i > 0 \ . \tag{7.11}$$

Then every $x \in R^n$ that satisfies (7.1), (7.2) and (7.3), also satisfies

$$\sum_{j \in J} \beta_j x_j \geq 1 \ , \tag{7.6$'$}$$

where

$$\beta_j = \begin{cases} \min\limits_{m \in M} \max\limits_{i \in Q} \dfrac{\sigma^i a_j^i + m_i}{\sigma^i a_0^i} \ , & j \in J_1 \\ \max\limits_{i \in Q} \dfrac{\sigma^i a_j^i}{\sigma^i a_0^i} \ , & j \in J_2 \end{cases} \ . \tag{7.7$'$}$$

<u>Proof</u>. Given any σ^i, $i \in Q$, satisfying (7.11), if we apply Theorem 7.1 by setting $\theta^i = (\sigma^i/\sigma^i a_0^i)$, $i \in Q$, in (7.7) and (7.8), we obtain the cut (7.6$'$), with β_j defined by (7.7$'$), $j \in J$. ||

Note that the cut-strengthening procedure of Theorem 7.1 requires, in order to be applicable, the existence of lower bounds on each component of $A^i x$, $\forall i \in Q$. This is a genuine restriction, but one that is satisfied in many practical instances. Thus, if x is the vector of nonbasic variables associated with a given basis, assuming that $A^i x$ is bounded below for each $i \in Q$ amounts to assuming that the basic variables are bounded below and/or above. In the case of a 0-1 program, for instance, such bounds not only exist but are quite tight.

Example 7.1. Consider again the mixed-integer program of example 3.1 (taken from [35]), and assume this time that x_1 and x_2 are 0-1 variables rather than just integer constrained, i.e., let the constraint set of the problem be given by

$$x_1 = .2 + .4(-x_3) + 1.3(-x_4) - .01(-x_5) + .07(-x_6)$$
$$x_2 = .9 - .3(-x_3) + .4(-x_4) - .04(-x_5) + .1(-x_6)$$
$$x_j \geq 0,\quad j = 1,\ldots,6;\quad x_j = 0 \text{ or } 1,\quad j = 1,2;\quad x_j \text{ integer},\quad j = 3,4.$$

This change does not affect the Gomory cuts or the cuts obtainable from extreme valid inequalities for the group problem, which remain the same as listed in example 3.1.

Now let us derive a cut, strengthened by the above procedure, from the disjunction

$$\left\{\begin{matrix} x_1 \geq 0 \\ -x_2 \geq 0 \end{matrix}\right\} \vee \{x_2 \geq 1\} .$$

Since x_1, $-x_2$ and x_2 are bounded below by 0, -1 and 0 respectively, we have

$$a_0^1 = \begin{pmatrix} -.2 \\ .9 \end{pmatrix}, \quad b_0^1 = \begin{pmatrix} -.2 \\ -.1 \end{pmatrix}; \quad a_0^2 = .1, \quad b_0^2 = -.9 .$$

Applying Corollary 7.1.1, we choose $\sigma^1 = (4,1)$, $\sigma^2 = 1$, which is easily seen to satisfy (7.11). Since Q has only 2 elements, the set M of (7.5).

becomes

$$M = \{m = (m_1, m_2) \mid m_1 + m_2 \geq 0;\ m_1,\ m_2 \text{ integer}\}$$

and, since at the optimum we may assume equality,

$$M = \{m = (m_1, -m_1) \mid m_1 \text{ integer}\} \quad .$$

The coefficients defined by (7.7′) then become

$$\beta_3 = \min_{m_1 \text{ integer}} \max \left\{ \frac{4 \times (-.4) + 1 \times (-.3) + m_1}{4 \times (-.2) + 1 \times .9} , \frac{1 \times .3 - m_1}{1 \times .1} \right\} = -7 \text{ (with } m_1 = 1\text{)}$$

$$\beta_4 = \min_{m_1 \text{ integer}} \max \left\{ \frac{4 \times (-1.3) + 1 \times .4 + m_1}{4 \times (-.2) + 1 \times .9} , \frac{1 \times (-.4) - m_1}{1 \times .1} \right\} = -24 \text{ (with } m_1 = 2\text{)}$$

$$\beta_5 = \max \left\{ \frac{4 \times .01 + 1 \times (-.04)}{4 \times (-.2) + 1 \times .9} , \frac{1 \times .04}{1 \times .1} \right\} = .4$$

$$\beta_6 = \max \left\{ \frac{4 \times (-.07) + 1 \times .1}{4 \times (-.2) + 1 \times .9} , \frac{1 \times (-.1)}{1 \times .1} \right\} = -1$$

and the cut is

$$-7x_3 - 24x_4 + .4x_5 - x_6 \geq 1 \quad ,$$

which has a smaller coefficient for x_4 (and hence is stronger) than the cut derived in example 3.1.

In the above example, the integers m_i were chosen by inspection. Derivation of an optimal set of m_i requires the solution of a special type of optimization problem. Two efficient algorithms are available [14] for doing this when the multipliers σ^i are fixed. Overall optimization would of course require the simultaneous choice of the σ^i and the m_i, but a good method for doing that is not yet available.

The following algorithm, which is one of the two procedures given in [14], works with fixed σ^i, $i \in Q$. It first finds optimal noninteger values for the m_i, $i \in Q$, and rounds them down to produce an initial set of integer values. The optimal integer values, and the corresponding value of β_j, are then found by applying an iterative step k times, where $k \leq |Q| - 1$, $|Q|$ being the number of terms in the disjunction from which the cut is derived.

<u>Algorithm for calculating β_j, $j \in J_1$, of (7.7′)</u>

Denote

$$\alpha_i = \sigma^i a^i_j \, , \qquad \lambda_i = (\sigma^i a^i_0)^{-1} \, , \tag{7.12}$$

and

$$\gamma = \sum_{i \in Q} \alpha_i \Big/ \sum_{i \in Q} \frac{1}{\lambda_i} \, . \tag{7.13}$$

Calculate

$$m_i^* = \frac{\gamma}{\lambda_i} - \alpha_i \, , \qquad i \in Q \, , \tag{7.14}$$

set $m_i = [m_i^*]$, $i \in Q$, define $k = - \sum_{i \in Q} [m_i^*]$, and apply k times the following

<u>Iterative Step</u>. Find

$$\lambda_s(\alpha_s + m_s + 1) = \min_{i \in Q} \lambda_i(\alpha_i + m_i + 1)$$

and set

$$m_s \leftarrow m_s + 1 \, , \qquad m_i \leftarrow m_i \, , \qquad i \in Q \setminus \{s\} \, .$$

This algorithm was shown in [14] to find an optimal set of m_i (and the associated value of β_j) in k steps, where $k = -\sum_{i \in Q} [m_i^*] \leq |Q| - 1$.

Example 7.2. Consider the integer program with the constraint set

$$x_1 = \frac{1}{6} + \frac{7}{6}(-x_5) - \frac{2}{6}(-x_6) + \frac{5}{6}(-x_7)$$

$$x_2 = \frac{2}{6} + \frac{1}{6}(-x_5) + \frac{1}{6}(-x_6) - \frac{1}{6}(-x_7)$$

$$x_3 = \frac{3}{6} - \frac{2}{6}(-x_5) + \frac{4}{6}(-x_6) - \frac{1}{6}(-x_7)$$

$$x_4 = \frac{1}{6} + \frac{4}{6}(-x_5) + \frac{5}{6}(-x_6) - \frac{1}{6}(-x_7)$$

$$x_1 + x_2 + x_3 + x_4 \geq 1$$

$$x_j = 0 \text{ or } 1,\ j = 1,\dots,4; \quad x_j \geq 0 \text{ integer},\ j = 5,6,7.$$

We wish to generate a strengthened cut from the disjunction

$$x_1 \geq 1 \ \vee\ x_2 \geq 1 \ \vee\ x_3 \geq 1 \ \vee\ x_4 \geq 1 .$$

If we apply Theorem 3.1 without strengthening and choose $\theta^i = (1 - a_{i0})^{-1}$, $i = 1,2,3,4$, we obtain the cut

$$\frac{2}{3}x_5 + \frac{2}{5}x_6 + \frac{1}{3}x_7 \geq 1 ,$$

whose j-th coefficient is

$$\alpha_j = \max_{i \in \{1,2,3,4\}} \frac{-a_{ij}}{1 - a_{i0}} .$$

To apply the strengthening procedure, we note that each x_j, $j = 1,2,3,4$ is bounded below by 0. Using $\sigma^i = 1$ (which satisfies (6.11) since $\sigma^i(a_0^i - b_0^i) = \sigma^i[1 - a_{i0} - (-a_{i0})] = 1$, $i = 1,2,3,4$, and $\sigma^i a_0^i = \sigma^i(1 - a_{i0}) > 0$, we obtain

$$\beta_5 = \min_{m \in M} \max \left\{ \frac{6}{5}\left(-\frac{7}{6}+m_1\right), \frac{6}{4}\left(-\frac{1}{6}+m_2\right), \frac{6}{3}\left(\frac{2}{6}+m_3\right), \frac{6}{5}\left(-\frac{4}{6}+m_4\right) \right\}$$

$$\beta_6 = \min_{m \in M} \max \left\{ \frac{6}{5}\left(\frac{2}{6}+m_1\right), \frac{6}{4}\left(-\frac{1}{6}+m_2\right), \frac{6}{3}\left(-\frac{4}{6}+m_3\right), \frac{6}{5}\left(-\frac{5}{6}+m_4\right) \right\}$$

$$\beta_7 = \min_{m \in M} \max \left\{ \frac{6}{5}\left(-\frac{5}{6}+m_1\right), \frac{6}{4}\left(\frac{1}{6}+m_2\right), \frac{6}{3}\left(\frac{1}{6}+m_3\right), \frac{6}{5}\left(\frac{1}{6}+m_4\right) \right\} .$$

Next we apply the above Algorithm for calculating β_j:

— For $j = 5$: $\gamma = -\frac{10}{17}$; $m_1^* = \frac{23}{34}$, $m_2^* = -\frac{23}{102}$, $m_3^* = -\frac{32}{51}$, $m_4^* = \frac{11}{51}$.

Thus our starting values are $[m_1^*] = 0$, $[m_2^*] = -1$, $[m_3^*] = -1$, $[m_4^*] = 0$.

Since $k = -(-1) - (-1) = 2$, the Iterative step is applied twice:

1. $\min \{-\frac{1}{5}, -\frac{1}{4}, \frac{2}{3}, \frac{2}{5}\} = -\frac{1}{4}$, $s = 2$; $m_1 = 0$, $m_2 = -1+1 = 0$, $m_3 = -1$, $m_4 = 0$.

2. $\min \{-\frac{1}{5}, \frac{5}{4}, \frac{2}{3}, \frac{2}{5}\} = -\frac{1}{5}$, $s = 1$; $m_1 = 1$, $m_2 = 0$, $m_3 = -1$, $m_4 = 0$.

These are the optimal m_i, and

$$\beta_5 = \max \{-\frac{1}{5}, -\frac{1}{4}, -\frac{4}{3}, -\frac{4}{5}\} = -\frac{1}{5} .$$

— For $j = 6$; $\gamma = -\frac{8}{17}$, $[m_1^*] = -1$, $[m_2^*] = -1$, $[m_3^*] = 0$, $[m_4^*] = 0$; $k = 2$.

1. $\min \{ \frac{2}{5}, -\frac{1}{4}, \frac{2}{3}, \frac{1}{5} \} = -\frac{1}{4}$, $s = 2$; $m_1 = -1$, $m_2 = 0$, $m_3 = 0$, $m_4 = 0$

2. $\min \{ \frac{2}{5}, \frac{5}{4}, \frac{2}{3}, \frac{1}{5} \} = \frac{1}{5}$, $s = 4$; $m = -1$, $m = 0$, $m = 0$, $m = 1$.

$$\beta_6 = \max \{-\frac{4}{5}, -\frac{1}{4}, -\frac{4}{3}, \frac{1}{5} \} = \frac{1}{5} .$$

— For $j = 7$: $\gamma = -\frac{2}{17}$, $[m_1^*] = 0$, $[m_2^*] = -1$, $[m_3^*] = -1$, $[m_4^*] = -1$; $k = 3$.

1. $\min \{ \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{5} \} = \frac{1}{5}$, $s = 1$; $m_1 = 1$, $m_2 = -1$, $m_3 = -1$, $m_4 = -1$;

2. $\min\{\frac{7}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{5}\} = \frac{1}{5}$, $s = 4$; $m_1 = 1$, $m_2 = -1$, $m_3 = -1$, $m_4 = 0$;

3. $\min\{\frac{7}{5}, \frac{1}{4}, \frac{1}{3}, \frac{7}{5}\} = \frac{1}{4}$, $s = 2$; $m_1 = 1$, $m_2 = 0$, $m_3 = -1$, $m_4 = 0$.

$\beta_7 = \max\{\frac{1}{5}, \frac{1}{4}, -\frac{5}{3}, \frac{1}{5}\} = \frac{1}{4}$.

Thus the strengthened cut is

$$-\frac{1}{5}x_5 + \frac{1}{5}x_6 + \frac{1}{4}x_7 \geq 1 .$$

The frequently occurring situation, when $|Q| = 2$, deserves special mention. In this case the coefficients β_j, $j \in J_1$, are given by

$$\beta_j = \min\{\lambda_1(\alpha_1 + \langle m_0^* \rangle), \lambda_2(\alpha_2 - [m_0^*])\} , \tag{7.15}$$

where

$$m_0^* = \frac{\lambda_2\alpha_2 - \lambda_1\alpha_1}{\lambda_1 + \lambda_2} , \tag{7.16}$$

with α_i, λ_i, $i = 1,2$, defined by (7.12), and $\langle m_0^* \rangle$ = the smallest integer $\geq m_0^*$. The optimal value of $m_1 = -m_2$ is either $\langle m_0^* \rangle$ or $[m_0^*]$, according to whether the minimum in (7.15) is attained for the first or the second term.

The strengthening procedure discussed in this section produces the seemingly paradoxical situation that weakening a disjunction by adding a new term to it, may result in a strengthening of the cut derived from the disjunction; or, conversely, dropping a term from a disjunction may lead to a weakening of the inequality derived from the disjunction. For instance, if the disjunction used in Example 7.2 is replaced by the stronger one

$$x_1 \geq 1 \quad \vee \quad x_2 \geq 1 \quad \vee \quad x_3 \geq 1 ,$$

then the cut obtained by the strengthening procedure is

$$-\frac{1}{5}x_5 + \frac{2}{5}x_6 + \frac{1}{4}x_7 \geq 1 ,$$

which is weaker than the cut of the example, since the coefficient of x_6 is 2/5 instead of 1/5. The explanation of this strange phenomenon is to be sought in the fact that the strengthening procedure uses the lower bounds on each term of the disjunction. In Example 7.2, besides the disjunction $x_1 \geq 1 \vee x_2 \geq 1 \vee x_3 \geq 1 \vee x_4 \geq 1$, the procedure also uses the information that $x_i \geq 0$, $i = 1,2,3,4$. When the above disjunction is strengthened by omitting the term $x_4 \geq 1$, then the procedure does not any more use the information that $x_4 \geq 0$.

8. Some Frequently Occurring Disjunctions

As mentioned earlier, one of the main advantages of the disjunctive programming approach is that it can make full use of the special structure inherent in many combinatorial problems. In [8], [9] cutting planes are derived from the logical conditions of the set partitioning problem, the linear complementarity problem, the two forms of representation for nonconvex separable programs, etc. More general complementarity problems are discussed in [33], [34]. Here we illustrate the procedure on the frequently occurring logical condition (where $x_i \geq 0$ integer, $i \in Q$)

$$(8.1) \qquad \sum_{i \in Q} x_i = 1,$$

often called a multiple choice constraint.

If all the problem constraints are of this form, we have a set partitioning problem. But the cut that we derive uses only one condition (8.1), so it applies to arbitrary integer programs with at least one equation (8.1). It also applies to mixed-integer programs, provided the variables x_i, $i \in Q$, are integer-constrained.

Let I and J be the index sets for the basic and nonbasic variables in a basic feasible noninteger solution of the form

$$x_i = a_{i0} + \sum_{j\in J} a_{ij}(-x_j) , \quad i\in I . \tag{8.2}$$

In [8], [9], several cutting planes are derived from the disjunction

$$\bigvee_{i\in Q} \begin{pmatrix} x_i = 1 \\ x_h = 0, \ \forall h \in Q\setminus\{i\} \end{pmatrix} .$$

Here we describe another cut, which in most cases turns out to be stronger than those mentioned above. It is derived from the disjunction

$$\sum_{i\in Q_1} x_i = 0 \ \vee \ \sum_{i\in Q_2} x_i = 0 , \tag{8.3}$$

clearly valid for any partition (Q_1,Q_2) of Q in the sense of being satisfied by every integer x satisfying (8.1).

Denoting

$$\beta_j^k = \sum_{i\in I\cap Q_k} a_{ij} , \qquad k = 1,2; \ j\in J\cup\{0\} , \tag{8.4}$$

(8.3) can be written as

$$\left\{\begin{array}{l} \sum_{j\in J} \beta_j^1 x_j \ge \beta_0^1 \\ \\ \sum_{j\in J\cap Q_1} x_j = 0 \end{array}\right\} \vee \left\{\begin{array}{l} \sum_{j\in J} \beta_j^2 x_j \ge \beta_0^2 \\ \\ \sum_{j\in J\cap Q_2} x_j = 0 \end{array}\right\}$$

which implies the disjunction

$$\sum_{j\in J\setminus Q_1} \beta_j^1 x_j \ge \beta_0^1 \ \vee \ \sum_{j\in J\setminus Q_2} \beta_j^2 x_j \ge \beta_0^2 , \tag{8.5}$$

with $\beta_0^k > 0$, $k = 1,2$. Note that once the sets $I \cap Q_k$, $k = 1,2$, are chosen the sets $J \cap Q_1$ and $J \cap Q_2$ can be "optimized," in the sense of putting an index $j \in J \cap Q$ into Q_1 if $(\beta_j^1/\beta_0^1) \geq (\beta_j^2/\beta_0^2)$, and in Q_2 otherwise. Using this device while applying Theorem 3.1 to (8.5) with multipliers $\theta^k = 1/\beta_0^k$, $k = 1,2$, we obtain the cut

$$\Sigma\, \beta_j x_j \geq 1 \;, \tag{8.6}$$

with coefficients

$$\beta_j = \begin{cases} \max\left\{\dfrac{\beta_j^1}{\beta_0^1}, \dfrac{\beta_j^2}{\beta_0^2}\right\}, & j \in J\setminus Q \\[2ex] \max\left\{0\;,\; \min\left\{\dfrac{\beta_j^1}{\beta_0^1}, \dfrac{\beta_j^2}{\beta_0^2}\right\}\right\}, & j \in J \cap Q \end{cases} \tag{8.7}$$

We now apply the strengthening procedure of section 7 to the coefficients β_j, $J\setminus Q$ (the coefficients indexed by $J \cap Q$ can usually not be further strengthened). This of course assumes that all x_j, $j \in J\setminus Q$, are integer constrained. A lower bound on $\sum_{j\in J\setminus Q_k} \beta_j^k x_j$ is $\beta_0^k - 1$, for $k = 1,2$, since

$$\sum_{i\in I\cap Q_k} x_i = \beta_0^k + \sum_{j\in J\setminus Q_k} \beta_j^k(-x_j) \leq 1 \;, \quad k = 1,2.$$

The multipliers $\sigma^k = 1$, $k = 1,2$, satisfy condition (7.11) of Corollary 7.1.1, since

$$\sigma^k[\beta_0^k - (\beta_0^k - 1)] = 1, \quad \sigma^k\beta_0^k > 0, \quad k = 1,2,$$

and thus the j-th coefficient of the strengthened cut becomes (Corollary 7.1.1)

$$\beta_j = \min_{m\in M}\ \max_{k\in\{1,2\}} \left\{ \frac{\beta_j^k + m_k}{\beta_0^k} \right\} .$$

with M defined by (7.5), with $|Q| = 2$. Applying the closed form solution (7.15) to the minimization problem involved in calculating β_j (in the special case of a disjunction with only two terms), we obtain

$$\alpha_1 = \sigma^k\beta_j^k = \beta_j^k , \quad \lambda_k = (\sigma^k\beta_0^k)^{-1} = \frac{1}{\beta_0^k} , \quad k = 1,2,$$

and hence

$$\beta_j = \min \left\{ \frac{\beta_j^1 + \langle m_0^*\rangle}{\beta_0^1} , \frac{\beta_j^2 - [m_0^*]}{\beta_0^2} \right\}$$

where

(8.8)
$$m_0^* = \frac{\beta_j^2\beta_0^1 - \beta_j^1\beta_0^2}{\beta_0^1 + \beta_0^2} .$$

We have thus proved

<u>Theorem 8.1</u>. If (8.2) is a basic feasible noninteger solution of the linear program associated with an integer program whose variables have to satisfy (8.1), then for any partition $(I\cap Q_1, I\cap Q_2)$ of the set $I\cap Q$, the inequality (8.6) is a valid cut, with coefficients

(8.9)
$$\beta_j = \begin{cases} \max\left\{0 , \min\left\{\dfrac{\beta_j^1}{\beta_0^1} , \dfrac{\beta_j^2}{\beta_0^2}\right\}\right\} , & j\in J\cap Q \\[2ex] \min\left\{\dfrac{\beta_j^1 + \langle m_0^*\rangle}{\beta_0^1} , \dfrac{\beta_j^2 - [m_0^*]}{\beta_0^2}\right\} , & j\in J\setminus Q \end{cases}$$

where the β_j^k, $k = 1,2$, $j \in J$, are defined by (8.4) and m_0^* is given by (8.8).

We illustrate this cut on a set partitioning problem, which is a special case of the Theorem.

Example 8.1. Consider the set partitioning problem whose cost vector is $c = (5,4,3,2,2,3,1,1,1,0)$ and whose coefficient matrix is given in Table 8.1.

	1	2	3	4	5	6	7	8	9	10
1	1					1	1		1	
2	1	1								1
3		1	1			1	1	1		
4		1	1	1			1			
5			1	1	1	1		1	1	

Table 8.1

The linear programming optimum is obtained for $x_4 = x_8 = x_9 = \frac{1}{3}$, $x_7 = \frac{2}{3}$, $x_{10} = 1$, and $x_j = 0$ for all other j. The associated system (8.2) is shown in the form of a simplex tableau in Table 8.2 (artificial variables have been removed).

	1	$-x_1$	$-x_6$	$-x_3$	$-x_5$	$-x_2$
x_0	-2	5	2	1	1	3
x_{10}	1	1	0	0	0	1
x_8	$\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
x_7	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$
x_4	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$
x_9	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$

Table 8.2

We choose the disjunction corresponding to row 5 of the matrix shown in Table 8.1, which is of the form (8.1), with $Q = \{3,4,5,6,8,9\}$.

We define $I \cap Q_1 = \{4,8\}$, $I \cap Q_2 = \{9\}$, and we have

	j = 1	j = 6	j = 3	j = 5	j = 2	j = 0
β_j^1	$-\frac{2}{3}$	0	$\frac{4}{3}$	0	$\frac{2}{3}$	$\frac{2}{3}$
β_j^2	$\frac{2}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$\frac{1}{3}$

Since $J \cap Q = \{3,5,6\}$, we need the values m_0^* for $j \in J \setminus Q = \{1,2\}$. They are $m_0^*(1) = \frac{2}{3}$, $m_0^*(2) = -\frac{2}{3}$. Hence

$$\beta_1 = \min\left\{\frac{-\frac{2}{3}+1}{\frac{2}{3}}, \frac{\frac{2}{3}-0}{\frac{1}{3}}\right\} = \frac{1}{2}$$

$$\beta_6 = \min\left\{0, \max\left\{\frac{0}{\frac{2}{3}}, \frac{\frac{2}{3}}{\frac{1}{3}}\right\}\right\} = 0$$

$$\beta_3 = \min\left\{0, \max\left\{\frac{\frac{4}{3}}{\frac{2}{3}}, \frac{-\frac{1}{3}}{\frac{1}{3}}\right\}\right\} = 0$$

$$\beta_5 = \min\left\{0\ ,\ \max\left\{\frac{0}{\frac{2}{3}}\ ,\ \frac{-\frac{2}{3}}{\frac{1}{3}}\right\}\right\} = 0$$

$$\beta_2 = \min\left\{\frac{\frac{2}{3}+0}{\frac{2}{3}}\ ,\ \frac{-\frac{2}{3}-(-1)}{\frac{1}{3}}\right\} = 1$$

and we obtain the cut

$$\frac{1}{2}x_1 + x_2 \geq 1$$

or

$$\begin{array}{c|cccccc|}
 & 1 & -x_1 & -x_6 & -x_3 & -x_5 & -x_2 \\
\hline
s & -1 & -\frac{1}{2} & 0 & 0 & 0 & -1 \\
\hline
\end{array}$$

which is considerably stronger than the traditional cuts that one can derive from Table 8.2, and it actually implies that the nonbasic variable x_2 has to be 1 in any integer solution.

Dual cutting plane methods have been found reasonably successful on set partitioning problems. Using stronger cuts can only enhance the efficiency of such methods, since the computational cost of the cut (8.9) is quite modest.

9. Combining Cutting Planes with Branch and Bound

The disjunctive programming approach offers various ways of combining branch and bound with cutting planes, some of which are currently the object of computational testing. Here we discuss one feature which seems to us crucial.

For any problem P, let v(P) denote the value of P (i.e., of an optimal solution to P).

Suppose we are using branch and bound to solve a mixed-integer 0-1 program P, stated as a maximization problem. If $\{P_i\}_{i\epsilon Q}$ is the set of active subproblems (active nodes of the search tree) at a given stage of the procedure and $\bar{v}(P_i)$ is the available upper bound on $v(P_i)$ (assume, for the sake of simplicity, that $\bar{v}(P_i) = v(LP_i)$, where LP_i is the linear program associated with P_i), then $\max_{i\epsilon Q} \bar{v}(P_i)$ is an upper bound on v(P). Also, $\underline{v}(P)$, the value of the best integer solution found up to the stage we are considering, is of course a lower bound on v(P); i.e., at any stage in the procedure ,

$$(9.1) \qquad \underline{v}(P) \leq v(P) \leq \max_{i\epsilon Q} \bar{v}(P_i) .$$

Hence the importance of finding good bounds for both sides of (9.1).

It is a crucial feature of the approach reviewed here that it can be used to derive a cutting plane from the optimal simplex tableaus associated with the subproblems LP_i, $i\epsilon Q$, which provides an upper bound on v(P) at least as good as, and often better than $\max_{i\epsilon Q} \bar{v}(P_i)$.

Let the linear program LP associated with the mixed-integer 0-1 program P have an optimal solution of the form

$$(9.2) \qquad x_h = a_{h0} + \sum_{j\epsilon J} a_{hj}(-x_j) , \qquad h \epsilon I \cup \{0\}$$

where I and J are the index sets for the basic and nonbasic variables respectively, and let I_1 and J_1 be the respective index sets for the integer constrained variables. Here $a_{h0} \geq 0$, $h \in I$, and $a_{h0} \leq 1$, $h \in I_1$. Further, since P is a maximization problem and the solution (9.2) is optimal, $a_{0j} \geq 0$, $j \in J$.

Now let $\{P_k\}_{k \in Q}$ be the set of active subproblems, and for $k \in Q$, let the optimal solution to LP_k, the linear program associated with P_k, be of the form

$$(9.3)_k \qquad x_h = a^k_{h0} + \sum_{j \in J^k} a^k_{hj}(-x_j) , \qquad h \in I^k \cup \{0\} ,$$

where I^k, J^k are defined with respect to LP_k the same way as I, J with respect to LP. Again $a^k_{0j} \geq 0$, $\forall j \in J^k$, since each LP_k is a maximization problem.

In order to derive a valid cutting plane from $(9.3)_k$, $k \in Q$, we view the branching process as the imposition of the disjunction

$$(9.4) \qquad \bigvee_{k \in Q} \begin{pmatrix} D^k x \geq d^k_0 \\ Ax \geq b \\ x \geq 0 \end{pmatrix} .$$

where $Ax \geq b$ stands for the system

$$\sum_{j \in J} (-a_{hj}) x_j \geq -a_{h0} , \qquad h \in I$$

$$\sum_{j \in J} a_{hj} x_j \geq a_{h0} - 1 , \qquad h \in I_1 ,$$

expressing the conditions $x_h \geq 0$, $h \in I$, $x_h \leq 1$, $h \in I_1$, while each $D^k x \geq d^k_0$ is composed of inequalities of the form

$$\sum_{j\in J} a_{ij}x_j \geq a_{i0}$$

or

$$\sum_{j\in J} (-a_{ij})x_j \geq 1 - a_{i0} ,$$

corresponding to the conditions $x_i \leq 0$ or $x_i \geq 1$ whose totality, together with $Ax \geq b$, $x \geq 0$, defines P_k.

Now consider the cut derived from (9.4) on the basis of Theorem 3.1, with the optimal dual variables (obtained by solving LP_k) used as the multipliers θ^k, $k \in Q$. If for $k \in Q$ we denote by (u^k, v^k) the optimal dual vector associated with the k-th term of (9.4) and

(9.5) $$\alpha^k = u^k D^k + v^k A , \qquad \alpha_0^k = u^k d_0^k + v^k b ,$$

and if $\alpha_0^k > 0$, $k \in Q$, then according to Theorem 3.1, the inequality

(9.6) $$\sum_{j\in J} (\max_{k\in Q} \frac{\alpha_j^k}{\alpha_0^k}) x_j \geq 1$$

is satisfied by every x satisfying (9.4), i.e., by every feasible integer solution. The condition $\alpha_0^k > 0$, $k \in Q$, amounts to requiring that $v(LP_k) < v(LP)$, i.e., that the "branching constraints" $D^k x \geq d_0^k$ force $v(LP\ )$ strictly below $v(LP)$, $\forall k \in Q$. This is a necessary and sufficient condition for the procedure discussed here to be applicable. Should the condition not be satisfied for some $k \in Q$, one can use a different objective function for LP_k than for the rest of the subproblems — but we omit discussing this case here. Note that, since $b \leq 0$, $v^k b \leq 0$, $\forall k \in Q$, and thus $\alpha_0^k > 0$ implies $u^k d_0^k > 0$, $\forall k \in Q$.

Since the multipliers (u^k, v^k) are optimal solutions to the linear programs dual to LP_k, $k \in Q$, they maximize the right hand side coefficient

α_0^k of each inequality $\alpha^k x \geq \alpha_0^k$ underlying the cut (9.6) subject to the condition that $\alpha_j^k \leq a_{0j}$, $\forall j \in J$.

We now proceed to strengthen the inequality (9.6) via the procedure of section 7. To do this, we have to derive lower bounds on $\alpha^k x$, $k \in Q$. We have

$$\begin{aligned} \alpha^k x - \alpha_0^k &= u^k(D^k x - d_0^k) + v^k(Ax - b) \\ &\geq u^k(D^k x - d_0^k) \geq -u^k e \end{aligned}$$

where $e = (1,\dots,1)$. The first inequality holds since $Ax - b \geq 0$ for all x satisfying (9.4), while the second one follows from the fact that each inequality of the system $D^k x - d_0^k \geq 0$ is either of the form $-x_i \geq 0$ or of the form $x_i - 1 \geq 0$, and in both cases -1 is a lower bound on the value of the left hand side. Thus

$$\alpha^k x \geq \alpha_0^k - u^k e, \qquad k \in Q \tag{9.7}$$

holds for every x satisfying (9.4). Note that $u^k d_0^k > 0$ implies $u^k \neq 0$ and hence $u^k \geq 0$ implies $u^k e > 0$, $k \in Q$.

We now apply Corollary 7.1.1 to the system

$$\begin{aligned} &\alpha^k x \geq \alpha_0^k - u^k e, \quad k \in Q, \\ &x \geq 0, \\ &\bigvee_{k \in Q} (\alpha^k x \geq \alpha_0^k), \\ &x_j \text{ integer}, \ j \in J_1. \end{aligned} \tag{9.8}$$

We choose $\sigma^k = 1/u^k e$, $k \in Q$, which satisfies condition (7.11) of the Corollary:

$$(1/u^k e)[\alpha_0^k - (\alpha_0^k - u^k e)] = 1 , \qquad (1/u^k e)\alpha_0^k > 0 .$$

The strengthened cut is then

$$\sum_{j\in J} \beta_j x_j \geq 1 , \tag{9.9}$$

with

$$\beta_j = \begin{cases} \min\limits_{m\in M} \max\limits_{k\in Q} \dfrac{(\alpha_j^k/u^k e) + m_k}{\alpha_0^k/u^k e} , & j \in J_1 \\[2ex] \max\limits_{k\in Q} \dfrac{\alpha_j^k}{\alpha_0^k} , & j \in J \setminus J_1 \end{cases} \tag{9.10}$$

where

$$M = \{m \in R^{|Q|} \mid \sum_{k\in Q} m_k \geq 0,\ m_k \text{ integer},\ k \in Q\} . \tag{9.11}$$

The values of α_j^k, α_0^k and $u^k e$ needed for computing the cut coefficients, are readily available from the cost row of the simplex tableaux associated with the optimal solutions to LP and LP_k, $k \in Q$. If the latter are represented in the form (9.2) and $(9.3)_k$ respectively, and if d_j^k and a_j denote the j-th column of D^k and A of (9.4), while S_k is the row index set for D^k, $k \in Q$, we have for all $k \in Q$

$$\begin{aligned} a_{0j}^k &= a_{0j} - u^k d_j^k - v^k a_j \\ &= a_{0j} - \alpha_j^k , \qquad && j \in J^k \cap J \end{aligned}$$

and

$$a_{0j}^k = 0 + u_j^k , \qquad j \in J^k \cap S_k ,$$

since the indices $j \in S_k$ correspond to the slack variables of the system

$-D^k x \leq -d_0^k$, whose costs are 0 (note that $S_k \cap J = \emptyset$ by definition). Further, for $j \in J \setminus J^k = J \cap I^k$ the reduced cost of x_j in LP_k is 0, hence for all $k \in Q$

$$\begin{aligned} 0 &= a_{0j} - u^k d_j^k - v^k a_j \\ &= a_{0j} - \alpha_j^k . \end{aligned}$$

Finally,

$$\begin{aligned} a_{00}^k &= a_{00} - u^k d_0^k - v^k b \\ &= a_{00} - \alpha_0^k , \qquad \forall k \in Q . \end{aligned}$$

From the above expressions we then have for $k \in Q$,

(9.12)
$$\alpha_j^k = \begin{cases} a_{0j} - a_{0j}^k , & j \in J \cap J^k \\ a_{0j} , & j \in J \setminus J^k \end{cases}$$

$$\alpha_0^k = a_{00} - a_{00}^k ,$$

and

(9.13)
$$u^k e = \sum_{i \in J^k \cap S_k} a_{0i}^k .$$

(since $u_i^k = 0$, $\forall i \in S_k \cap I^k$).

The representation $(9.3)_k$ of the optimal solution to LP_k assumes that the slack variable $j \in S_k$ of each "branching constraint" $x_i \leq 0$ or $x_i \geq 1$ that is tight at the optimum, is among the nonbasic variables with $a_{0j} > 0$. If one prefers instead to replace these slacks with the corresponding structural variables x_i and regard the latter as "fixed" at 0 or 1, and if F_k denotes the index set of the variables fixed in LP_k, the reduced costs

a_{0i}^k, $i \in J^k \cap F_k$ are then the same, except for their signs, as a_{0j}^k, $j \in J^k \cap S_k$, and the only change required in the expressions derived above is to replace (9.13) by

$$(9.13') \qquad u^k e = \sum_{i \in J^k \cap F_k} |a_{0i}^k| \, .$$

Of course, in order to calculate a cut of the type discussed here, one needs the reduced costs a_{0j} for both the free and the fixed variables.

We have thus proved the following result.

<u>Theorem 9.1</u>. If LP and LP_k, $k \in Q$, have optimal solutions of the form (9.2) and $(9.3)_k$ respectively, with $a_{00} > a_{00}^k$, $k \in Q$, then every feasible integer solution satisfies the inequality (9.9), with coefficients defined by (9.10), (9.11), (9.12) and (9.13) [or (9.13')].

In the special case when $|Q| = 2$ and LP_1, LP_2 are obtained from LP by imposing $x_i \leq 0$ and $x_i \geq 1$ respectively (for some $i \in I$ such that $0 < a_{i0} < 1$), the definition of β_j for $j \in J_1$ becomes

$$(9.10') \qquad \beta_j = \min \left\{ \frac{(\alpha_j^1/u_i^1) + \langle m_0^* \rangle}{\alpha_0^1/u_i^1} \, , \, \frac{(\alpha_j^2/u_i^2) - [m_0^*]}{\alpha_0^2/u_i^2} \right\}$$

with

$$(9.14) \qquad m_o^* = \frac{\alpha_j^2\alpha_0^1 - \alpha_j^1\alpha_0^2}{u_i^2\alpha_0^1 + u_i^1\alpha_0^2} \, .$$

We now state the property of the cut (9.9) mentioned at the beginning of this section.

<u>Corollary 9.1.1</u>. Adding the cut (9.9) to the constraints of LP and performing one dual simplex pivot in the cut row reduces the value of v(LP)

from a_{00} to $\bar{a}_{00}$ such that

(9.15) $$\bar{a}_{00} \leq \max_{k \in Q} a^k_{00} .$$

Proof. For each $k \in Q$, $a^k_{00} = a_{00} - \alpha^k_0$. Now suppose (9.1) is false, i.e., $\bar{a}_{00} > a^k_{00}$, $\forall k \in Q$. Then

$$\bar{a}_{00} = a_{00} - \min_{j \in J | \beta_j > 0} a_{0j}/\beta_j > a_{00} - \alpha^k_0 , \qquad \forall k \in Q ,$$

and hence for all $k \in Q$,

$$\alpha^k_0 > \min_{j \in J | \beta_j > 0} a_{0j}/\beta_j$$

(9.16) $$\geq \min_{j \in J | \alpha^s_j > 0} a_{0j}(\alpha^s_0/\alpha^s_j) ,$$

where

(9.17) $$\frac{\alpha^s_j}{\alpha^s_0} = \max_{h \in Q} \frac{\alpha^h_j}{\alpha^h_0} .$$

The second inequality in (9.16) holds since the cut (9.9) is a strengthened version of (9.6), in the sense that

$$\beta_j \leq \max_{h \in Q} \frac{\alpha^h_j}{\alpha^h_0} , \qquad j \in J .$$

Now suppose the minimum in the second inequality of (9.16) is attained for $j = t$. Since (9.16) holds for all $k \in Q$, we then have

$$\alpha^s_0 > a_{0t}(\alpha^s_0/\alpha^s_t)$$

or (since $\alpha_t^s > 0$, $\alpha_0^s > 0$), $\alpha_t^s > a_{0t}$. But this contradicts the relation $\alpha_t^s \le a_{0t}$ implied by (9.12). ||

Note that the Corollary remains true if the strengthened cut (9.9) is replaced by its weaker counterpart (9.6). In that case, however, (9.15) holds with equality. The remarkable fact about the property stated in the Corollary is that by using the strengthened cut (9.9) one often has (9.15) satisfied as strict inequality. More precisely, we have the following

Remark 9.1. (9.15) holds as strict inequality if $\beta_t < \alpha_t^s/\alpha_0^s$, where s is defined by (9.17), and t by

$$a_{0t}(\alpha_0^s/\alpha_t^s) = \min_{j\epsilon J|\alpha_j^s>0} a_{0j}(\alpha_0^s/\alpha_t^s) .$$

Note that for (9.15) to hold as strict inequality the pivot discussed in the Corollary need not occur on a "strengthened" cut coefficient. All that is needed, is that the coefficient on which the pivot would occur in case the unstrengthened cut were used, should be "strengthened" (i.e., reduced by the strengthening procedure).

The significance of the cut of Theorem 9.1 is that it concentrates in the form of a single inequality much of the information generated by the branch and bound procedure up to the point where it is derived, and thus makes it possible, if one so wishes, to start a new tree search while preserving a good deal of information about the earlier one.

Theorem 9.1 and its Corollary are stated for (pure or mixed) 0-1 programs; but the 0-1 property (as opposed to integrality) is only used for the derivation of the lower bounds on $D^k x \geq d_0^k$, $k \in Q$; hence it only involves those variables on which branching has occurred. The results are therefore valid for mixed-integer programs with some 0-1 variables, provided the strengthening procedure is only used on cuts derived from branching on 0-1 variables.

Example 9.1. Consider the problem in the variables $x_j \geq 0$, $j = 1,\ldots,6$; $x_j = 0$ or 1, $j = 1,4$; x_j integer, $j = 2,3$, whose linear programming relaxation has the optimal solution shown in Table 8.1.

	1	$-x_3$	$-x_4$	$-x_5$	$-x_6$
x_0	1.1	2.0	.2	.05	1.17
x_1	.2	.4	1.3	-.01	.07
x_2	.9	-.3	.4	-.04	.1

Table 9.1

If we solve the problem by branch and bound, using the rules of always selecting for branching (a) LP_k with the largest a_{00}^k, and (b) x_i with the largest max {up penalty, down penalty}, we generate the search tree shown in Fig. 9.1. The optimal solution is $x = (0,2,1,0,20,0)$, with value -1.9, found at node 6. To prove optimality, we had to generate two more nodes, i.e., the total number of nodes generated (apart from the starting node) is 8.

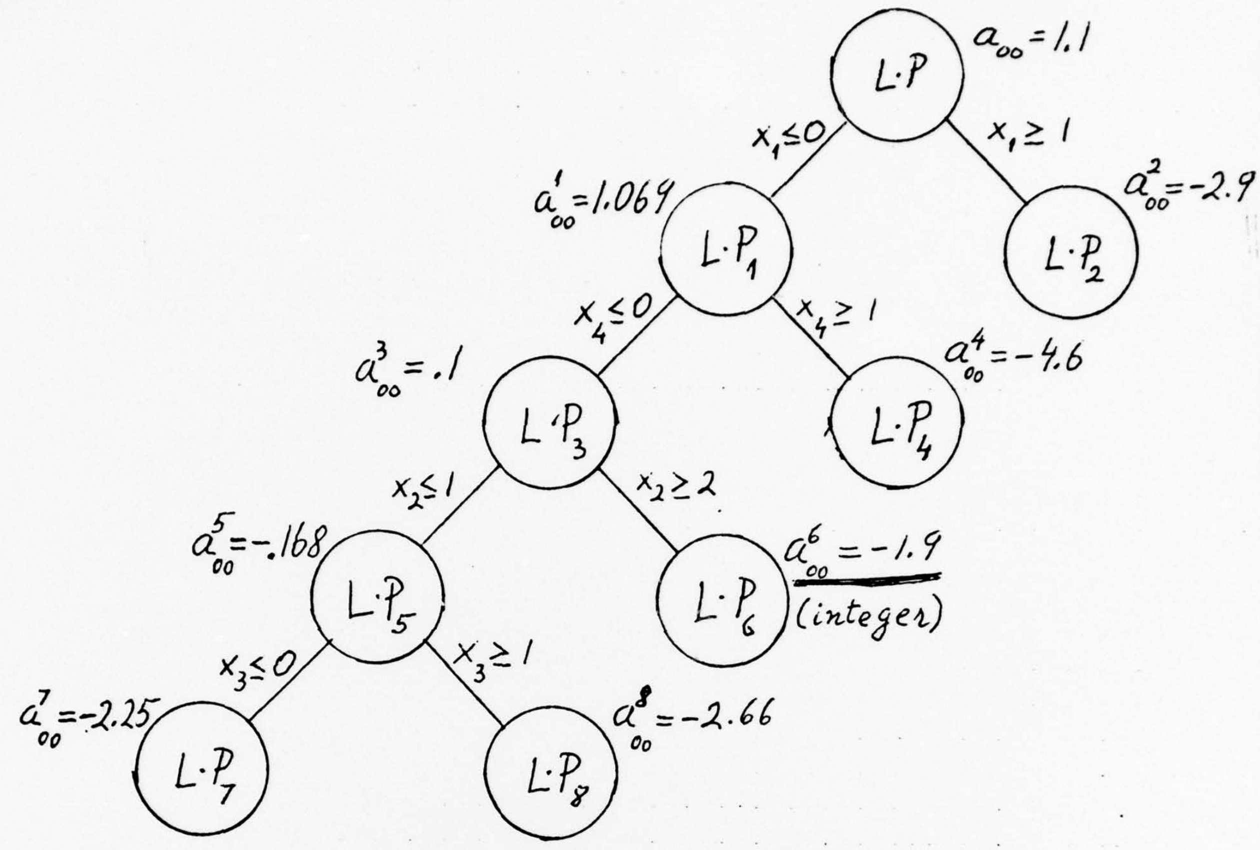


Fig. 9.1

Suppose now that after generating the first four nodes, we wish to use the available information to derive a cut. At that point there are three active nodes, associated with LP_k, for k = 2,3,4. The corresponding reduced cost coefficients a^k_{0j}, $j \in J^k$, are shown in Table 9.2. The slack variables of the "branching constraints" $x_1 \leq 0$, $x_1 \geq 1$, $x_4 \leq 0$, and $x_4 \geq 1$ are denoted by x_7, x_8, x_9 and x_{10} respectively.

k	2				3				4			
J^k	3	4	8	6	7	9	5	6	3	10	1	6
a^k_{0j}	4.	6.7	5.	1.52	5.	6.3	.1	.82	4.	6.7	5.	1.52
a^k_{00}	-2.9				.1				-4.6			

Table 9.2

The coefficients α_j^k, α_o^k and $u^k e$, extracted from Table 9.2 and the cost row of Table 9.1, are as follows:

__k = 2__: $\alpha_0^2 = a_{00} - a_{00}^2 = 4$; $u^2 e = u_1^2 = a_{08}^2 = 5$; $\alpha_3^2 = a_{03} - a_{03}^2 = -2$,

$\alpha_4^2 = a_{04} - a_{04}^2 = -6.5$, $\alpha_5^2 = a_{05} = .05$, $\alpha_6^2 = a_{06} - a_{06}^2 = -.35$;

__k = 3__: $\alpha_0^3 = 1$; $u^3 e = u_7^3 + u_9^3 = 11.3$; $\alpha_3^3 = 2$, $\alpha_4^3 = .2$, $\alpha_5^3 = -.05$, $\alpha_6^3 = .35$;

__k = 4__: $\alpha_0^4 = 5.7$; $u^4 e = u_{10}^4 = 6.7$; $\alpha_3^4 = -2$, $\alpha_4^4 = .2$, $\alpha_5^4 = .05$, $\alpha_6^4 = -.35$.

The coefficients of the strengthened cut are shown below, where $M = \{m \in R^3 \mid m_1 + m_2 + m_3 \geq 0,\ m_i \text{ integer},\ i = 1,2,3\}$.

$$\beta_3 = \min_{m \in M} \max \left\{ \frac{(-2/5) + m_1}{4/5}, \frac{(2/11.3) + m_2}{1/11.3}, \frac{(-2/6.7) + m_3}{5.7/6.7} \right\} = .75,$$

with m = (1,-1,0).

$$\beta_4 = \min_{m \in M} \max \left\{ \frac{(-6.5/5) + m_1}{4/5}, \frac{(.2/11.3) + m_2}{1/11.3}, \frac{(.2/6.7) + m_3}{5.7/6.7} \right\} = .035$$

with m = (1,-1,0)

$$\beta_5 = \max \left\{ \frac{.05}{4}, \frac{-.05}{1}, \frac{.05}{5.7} \right\} = .0125$$

$$\beta_6 = \max \left\{ \frac{-.35}{4}, \frac{.35}{1}, \frac{-.35}{5.7} \right\} = .35 .$$

Adding to the optimal tableau of LP the cut

$$.75\, x_3 + .035\, x_4 + .0125\, x_5 + .35\, x_6 \geq 1$$

produces Table 9.3 and the two pivots shown in Tables 9.3 and 9.4 produce the optimal Tableau 9.5. Thus no further branching is required.

	1	$-x_3$	$-x_4$	$-x_5$	$-x_6$
x_0	1.1	2.0	.2	.05	1.17
x_1	.2	.4	1.3	-.01	.07
x_2	.9	-.3	.4	-.04	.1
s	-1.0	$\boxed{-.75}$	-.035	-.0125	-.35

Table 9.3

	1	$-s$	$-x_4$	$-x_5$	$-x_6$
x_0	-1.57	2.67	.106	.01675	1.117
x_1	-.333	.533	1.281	$\boxed{-.01675}$	-.117
x_2	1.3	-.4	.414	-.035	.240
x_3	1.33	-1.33	.047	.01675	.467

Table 9.4

	1	$-s$	$-x_4$	$-x_5$	$-x_6$
x_0	-1.9	2.687	1.387	1.0	1.0
x_5	20.0	-31.8	-73.0	-60.0	7.0
x_2	2.0			-2.0	
x_3	1.0			1.0	

Table 9.5

10. Disjunctions from Conditional Bounds

In solving pure or mixed integer programs by branch and bound, the most widely used rule for breaking up the feasible set is to choose an integer-constrained variable x_i whose value a_{io} at the linear programming optimum is noninteger, and to impose the disjunction $(x_i \leq [a_{io}]) \vee (x_i \geq [a_{io}] + 1)$. It has been observed, however, that in the presence of multiple choice constraints, i.e., of constraints of the form

$$\sum_{i \in Q} x_i = 1 ,$$

it is more efficient to use a disjunction of the form

$$(\sum_{i \in Q_1} x_i = 0) \vee (\sum_{i \in Q_2} x_i = 0) ,$$

where $Q_1 \cup Q_2 = Q$, $Q_1 \cap Q_2 = \emptyset$, and Q_1 and Q_2 are about equal in size.

This is just one example of a situation where it is possible to branch so as to fix the values of several variables on each branch. The circumstance that makes this possible in the above instance is the presence of the rather tight multiple choice constraint. More generally, a tightly constrained feasible set makes it possible to derive disjunctions stronger than the usual dichotomy on a single variable. On the other hand, the feasible set of any integer program becomes more or less tightly constrained after the discovery of a "good" solution (in particular, of an optimal solution), provided that one restricts it to those solutions better than the current best. Such a "tightly constrained" state of the feasible set can be expressed in the form of an inequality

$$\pi x \leq \pi_o$$

with $\pi \geq 0$, and with a relatively small $\pi_o > 0$. One way of doing this, if the problem is of the form

(P) $\min\{cx | Ax \geq b,\ x \geq 0,\ x_j \text{ integer},\ j\epsilon N\}$,

with c integer, and if z_U is the value of the current best integer solution, is to find a set of multipliers u such that

$$uA \leq c,\ u \geq 0 ,$$

and define

$$\pi = c - uA,\ \pi_o = z_U - ub - 1.$$

Then multiplying $Ax \geq b$ by -u and adding the resulting inequality, $-uAx \leq -ub$, to $cx \leq z_U - 1$, yields the inequality

$$\pi x \leq \pi_o ,$$

satisfied by every feasible integer x such that $cx < z_U$. Here $\pi \geq 0$, $\pi_o > 0$, and the size of π_o depends on the gap between the upper bound z_U and the lower bound ub on the value of (P).

Now suppose we have such an inequality $\pi x \leq \pi_o$. Without loss of generality, we may assume that $\pi_j > 0$, $\forall$ j (by simply deleting the zero components). Then the following statement holds [12].

Theorem 10.1. Let $\pi \epsilon R^n$, $\pi_o \epsilon R$, $(\pi, \pi_o) > 0$, $N = \{1,\ldots,n\}$, and for $i = 1,\ldots,p$, $1 \leq p \leq n$, let $Q_i \subset N$, $Q_i \neq \emptyset$, with

$$\pi_{j(i)} = \min_{j\epsilon Q_i} \pi_j .$$

If the sets Q_i, $i = 1,\ldots,p$, satisfy the conditions

(10.1) $$\sum_{i | j\epsilon Q_i} \pi_{j(i)} \leq \pi_j ,\quad j\epsilon N,$$

and

$$(10.2) \qquad \sum_{i=1}^{p} \pi_{j(i)} > \pi_o ,$$

then every integer vector $x \geq 0$ such that $\pi x \leq \pi_o$, satisfies the disjunction

$$(10.3) \qquad \bigvee_{i=1}^{p} (x_j = 0, \; j \epsilon Q_i) .$$

<u>Proof</u>. Every integer $x \geq 0$ which violates (10.3) satisfies

$$(10.4) \qquad \sum_{j \epsilon Q_i} x_j \geq 1 \quad , \quad i = 1,\ldots,p.$$

Multiplying by $\pi_{j(i)}$ the i^{th} inequality of (10.4) and adding up the resulting inequalities, one obtains

$$(10.5) \qquad \sum_{j \epsilon \bigcup_{i=1}^{p} Q_i} \left(\sum_{i \mid j \epsilon Q_i} \pi_{j(i)} \right) x_j \geq \sum_{i=1}^{p} \pi_{j(i)} .$$

Further,

$$\sum_{j \epsilon N} \pi_j x_j \geq \sum_{j \epsilon \bigcup_{i=1}^{p} Q_i} \pi_j x_j$$

$$\geq \sum_{j \epsilon \bigcup_{i=1}^{p} Q_i} \left(\sum_{i \mid j \epsilon Q_i} \pi_{j(i)} \right) x_j \qquad \text{[from (10.1)]}$$

$$\geq \sum_{i=1}^{p} \pi_{j(i)} \qquad \text{[from (10.5)]}$$

$$> \pi_o \qquad \text{[from (10.2)]}.$$

<u>Q.E.D.</u>

One way of looking at Theorem 10.1 is as follows. Suppose the constraints of an integer program which include the inequality $\pi x \leq \pi_o$, were amended by the additional constraints (10.4).

From the proof of the Theorem, these inequalities give a lower bound on πx which exceeds π_o; this contradiction then produces the disjunction (10.3). Since the inequalities (10.4) are not actually part of the problem, we call the bound on πx derived from them a conditional bound, and the disjunction obtained from such bounds, a disjunction from conditional bounds.

Example 10.1. The inequality

$$9x_1+8x_2+8x_3+7x_4+7x_5+6x_6+6x_7+5x_8+5x_9+5x_{10}+4x_{11}+4x_{12}+3x_{13}+3x_{14}+3x_{15}+2x_{16} +2x_{17} \leq 10,$$

together with the condition $x \geq 0$, x_j integer $\forall j$, implies the disjunction

$$(x_j = 0,\ j = 1,2,3,4,5,6,7) \vee (x_j = 0,\ j = 1,8,9,10,11,12,13,14) \vee (x_j = 0,\ j = 2,3,8,9,10,15,16,17).$$

Indeed,

$$\pi_{j(1)} = \min\{9,8,8,7,7,6,6\} = 6,$$

$$\pi_{j(2)} = \min\{9,5,5,5,4,4,3,3\} = 3,$$

$$\pi_{j(3)} = \min\{8,8,5,5,5,3,2,2\} = 2,$$

and (10.1), (10.2) are satisfied, since $6 + 3 + 2 > 10$, while $6 + 3 \leq 9$ $(j = 1)$, $6 + 2 \leq 8$ $(j = 2,3)$, $6 \leq 7$ $(j = 4,5)$, $6 \leq 6$ $(j = 6,7)$, $3 + 2 \leq 5$ $(j = 8,9,10)$, $3 \leq 4$ $(j = 11,12)$, $3 \leq 3$ $(j = 13,14)$, $2 \leq 3$ $(j = 15)$, $2 \leq 2$ $(j = 16,17)$.

Next we outline a procedure [12] based on Theorem 10.1 for systematically generating disjunctions of the type (10.3) from an inequality $\pi x \leq \pi_o$, with $\pi_j > 0$, $\forall j$.

1. Choose some $S \subset N$ such that

$$\sum_{j \in S} \pi_j > \pi_o$$

but

$$\sum_{j \in T} \pi_j \leq \pi_o$$

for all $T \subset S$, $T \neq S$. Order $S = \{j(1),\ldots,j(p)\}$ according to decreasing values of $\pi_{j(i)}$ and go to 2.

2. Set

$$Q_1 = \{j \in N | \pi_j \geq \pi_{j(1)}\}$$

and define recursively

$$Q_i = \{j \in N | \pi_j \geq \pi_{j(i)} + \sum_{k=1}^{i-1} \pi_{j(k)} \delta_j^k\}, \quad i = 2,\ldots,p$$

where $\delta_j^k = 1$ if $j \in Q_k$, $\delta_j^k = 0$ otherwise. The sets Q_i, $i = 1,\ldots,p$, obtained in this way, satisfy (9.1), (9.2).

In the above example, $S = \{7,14,17\}$. If, on the other hand, one uses $S = \{5,12\}$ (which is also admissible, since $\pi_5 + \pi_{12} = 7 + 4 > 10$), one obtains the disjunction

$$(x_j = 0,\ j = 1,\ldots,5) \vee (x_j = 0,\ j = 6,\ldots,12).$$

A disjunction of the form (10.3) can be used to partition the feasible set into p subproblems, the k^{th} one of which is constrained by

$$\sum_{j \in Q_i} x_j \geq 1,\ i = 1,\ldots,k-1; \quad x_j = 0,\ \forall j \in Q_k .$$

Another way of using a disjunction of the form (10.3) is to derive cutting planes. This has been explored in the context of set covering problems [12] and found to yield good results. In particular, let $A = (a_{ij})$ be a 0-1 matrix and consider the set covering problem

$$\text{(SC)} \qquad \min\{cx \mid Ax \geq e,\ x_j = 0 \text{ or } 1,\ j \in N\}$$

where e is the vector of 1's of appropriate dimension. For every row i of A, let

$$N_i = \{j \in N \mid a_{ij} = 1\}.$$

Now suppose a prime cover (a basic feasible integer solution) $\bar{x}$ is known; then $z_U = c\bar{x}$ is an upper bound on the value of the optimum. If $\bar{u}$ is any feasible solution to the dual of the linear programming relaxation of (SC), i.e., any vector satisfying $\bar{u}A \leq c$, $\bar{u} \geq 0$, then setting $\pi = c - \bar{u}A$ and $\pi_o = z_U - \bar{u}e$ one obtains an inequality $\pi x \leq \pi_o$ which is satisfied by every integer solution better than $\bar{x}$, and which can therefore be used to derive a disjunction of form (10.3).

Suppose this is done, and a disjunction (10.3) is at hand, known to be satisfied by every feasible integer x better than the current best. Then for each $i \in \{1,\ldots,p\}$, one chooses a row h(i) of A, such that $N_{h(i)} \cap Q_i$ is "large" - or, conversely, $N_{h(i)} \setminus Q_i$ is "small." Clearly (and this is true for any choice of the indices h(i)), the disjunction (10.3) implies

$$\text{(10.6)} \qquad \bigvee_{i=1}^{p} \Big(\sum_{j \in N_{h(i)} \setminus Q_i} x_j \geq 1 \Big)$$

which in turn implies the inequality

$$\text{(10.7)} \qquad \sum_{j \in W} x_j \geq 1$$

where

$$W = \bigcup_{i=1}^{p} [N_{h(i)} \setminus Q_i].$$

The class of cutting planes (10.7) obtained in this way was shown in [12] to include as a proper subclass the Bellmore-Ratliff inequalities [16] derived from involutory bases. An all integer algorithm which uses these cutting planes was implemented and tested with good results on set covering problems with up to 900 variables (see [12] for details).

References

[1] J.A. Araoz Durand: "Polyhedral Neopolarities." Ph.D. Dissertation, University of Waterloo, November 1973.

[2] E. Balas: "Intersection Cuts - A New Type of Cutting Planes for Integer Programming." Operations Research, 19, 1971, p. 19-39.

[3] E. Balas: "Integer Programming and Convex Analysis: Intersection Cuts from Outer Polars." Mathematical Programming, 2, 1972, p. 330-382.

[4] E. Balas: "Ranking the Facets of the Octahedron." Discrete Mathematics, 2, 1972, p. 1-15.

[5] E. Balas: "Nonconvex Quadratic Programming via Generalized Polars." SIAM Journal of Applied Mathematics, 28, 1975, p. 335-349.

[6] E. Balas: "A Constraint-Activating Outer Polar Method for Pure and Mixed Integer 0-1 Programs," in Mathematical Programming in Theory and Practice, P.L. Hammer and G.T. Zoutendijk (editors), North-Holland, 1974, p. 275-310.

[7] E. Balas: "On the Use of Intersection Cuts in Branch and Bound." Paper presented at the 8th International Symposium on Mathematical Programming, Stanford, Calif., August 27-31, 1973.

[8] E. Balas: "Intersection Cuts from Disjunctive Constraints." MSRR No. 330, Carnegie-Mellon University. First draft, August 1973. Revised and expanded, February 1974.

[9] E. Balas: "Disjunctive Programming: Cutting Planes from Logical Conditions," in Nonlinear Programming 2, O.L. Mangasarian, R.R. Meyer and S.M. Robinson (editors), Academic Press, 1975, pp. 279-312.

[10] E. Balas: "Disjunctive Programming: Properties of the Convex Hull of Feasible Points." MSRR No. 348, Carnegie-Mellon University, July 1974.

[11] E. Balas: "A Note on Duality in Disjunctive Programming." Journal of Optimization Theory and Applications, 15, 1977, p.

[12] E. Balas: "Set Covering with Cutting Planes from Conditional Bounds." MSRR No. 399, Carnegie-Mellon University, July 1976.

[13] E. Balas, V.J. Bowman, F. Glover and D. Sommer: "An Intersection Cut from the Dual of the Unit Hypercube." Operations Research, 19, 1971, p. 40-44.

[14] E. Balas and R. Jeroslow: "Strengthening Cuts for Mixed Integer Programs." MSRR No. 359, Carnegie-Mellon University, February 1975.

[15] E. Balas and A. Zoltners: "Intersection Cuts from Outer Polars of Truncated Cubes." Naval Research Logistics Quarterly, 22, 1975, p. 477-496.

[16] M. Bellmore and H. D. Ratliff: "Set Covering and Involutory Bases." Management Science, 18, 1971, p. 194-206.

[17] C.E. Blair and R.G. Jeroslow: "A Converse for Disjunctive Constraints." MSRR No. 393, Carnegie-Mellon University, June 1976.

[18] C.A. Burdet: "Enumerative Inequalities in Integer Programming." Mathematical Programming, 2, 1972, p. 32-64.

[19] C.A. Burdet: "Polaroids: A New Tool in Nonconvex and Integer Programming." Naval Logistics Research Quarterly, 20, 1973, p. 13-24.

[20] D.R. Fulkerson: "Blocking Polyhedra," in Graph Theory and Its Applications, B. Harris (editor), Academic Press, 1972, p. 93-112.

[21] D.R. Fulkerson: "Anti-Blocking Polyhedra." Journal of Combinatorial Theory, 12, 1972, p. 50-71.

[22] D.R. Fulkerson: "Blocking and Anti-Blocking Pairs of Polyhedra." Mathematical Programming, 1, 1971, p. 168-194.

[23] F. Glover: "Convexity Cuts and Cut Search." Operations Research, 21, 1973, p. 123-134.

[24] F. Glover: "Convexity Cuts for Multiple Choice Problems." Discrete Mathematics, 6, 1973, p. 221-234.

[25] F. Glover: "Polyhedral Annexation in Mixed Integer Programming." MSRS 73-9, University of Colorado, August 1973. Revised and published as "Polyhedral Annexation in Mixed Integer and Combinatorial Programming," Mathematical Programming, 9, 1975, p. 161-188.

[26] F. Glover and D. Klingman: "The Generalized Lattice Point Problem." Operations Research, 21, 1973, p. 141-156.

[27] H.J. Greenberg and W.P. Pierskalla: "Stability Theorems for Infinitely Constrained Mathematical Programs." Journal of Optimization Theory and Applications, 16, 1975, p. 409-428.

[28] B. Grünbaum: Convex Polytopes, J. Wiley, 1967.

[29] Hoang Tuy: "Concave Programming Under Linear Constraints." (Russian), Doklady Akademii Narck SSSR, 1964. English translation in Soviet Mathematics, 1964, p. 1437-1440.

[30] R.G. Jeroslow: "The Principles of Cutting Plane Theory: Part I." Carnegie-Mellon University, February 1974.

[31] R.G. Jeroslow: "Cutting Planes for Relaxations of Integer Programs." MSRR No. 347, Carnegie-Mellon University, July 1974.

[32] R.G. Jeroslow: "Cutting Plane Theory: Disjunctive Methods." Annals of Discrete Mathematics, vol. 1: Studies in Integer Programming, 1977, p. 293-330.

[33] R.G. Jeroslow: "Cutting Planes for Complementarity Constraints." SIAM Journal on Control and Optimization, 16, 1978, p. 56-62.

[34] R.G. Jeroslow: "A Cutting Plane Game and Its Algorithms." CORE Discussion Paper 7724, June 1977.

[35] E.L. Johnson: "The Group Problem for Mixed Integer Programming." Mathematical Programming Study 2, 1974, p. 137-179.

[36] G. Owen: "Cutting Planes for Programs with Disjunctive Constraints." Journal of Optimization Theory and Applications, 11, 1973, p. 49-55.

[37] R.T. Rockafellar: Convex Analysis, Princeton University Press, 1970.

[38] J. Stoer and C. Witzgall: Convexity and Optimization in Finite Dimensions. I. Springer, 1970.

[39] J. Tind: "Blocking and Anti-Blocking Sets." Mathematical Programming, 6, 1974, p. 157-166.

[40] R.D. Young: "Hypercylindrically Deduced Cuts in Zero-One Integer Programs." Operations Research, 19, 1971, p. 1393-1405.

[41] E. Zemel: "On the Facial Structure of 0-1 Polytopes." Ph.D. Dissertation, GSIA, Carnegie-Mellon University, May 1976.

[42] P. Zwart: "Intersection Cuts for Separable Programming." Washington University, St. Louis, January 1972.

REPORT DOCUMENTATION PAGE | READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
Technical Report No. 415		

4. TITLE (and Subtitle) Disjunctive Programming	5. TYPE OF REPORT & PERIOD COVERED Technical Report August 1977
	6. PERFORMING ORG. REPORT NUMBER MSRR [redacted] 415
7. AUTHOR(s) Egon Balas	8. CONTRACT OR GRANT NUMBER(s) N0014-75-C-0621 MPS73-08534 A02 NSF
9. PERFORMING ORGANIZATION NAME AND ADDRESS Graduate School of Industrial Administration Carnegie-Mellon University Pittsburgh, Pennsylvania 15213	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 047-048
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 434) Arlington, Virginia 22217	12. REPORT DATE August 1977
	13. NUMBER OF PAGES 75
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Disjunctive programming, 0-1 programming, integer programming, logical conditions, polarity, reverse polars

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper reviews some recent developments in the convex analysis approach to integer programming. These developments are based on viewing integer programs as disjunctive programs, i.e., linear programs with disjunctive constraints, an approach which seems to be fruitful for 0-1 programming both theoretically and practically. On the theoretical side, it provides structural characterizations which offer new insights. On the practical side, it produces a variety of cutting planes with desirable properties and offers new ways of combining cutting planes with enumerative techniques.